# Exp Mail EV449562625US USAN 09/902,941



# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISI	HED (	JNDER THE PATENT COOPERATION TREATY (PCT)			
(51) International Patent Classification <sup>6</sup> :		(11) International Publication Number: WO 97/49417			
A61K 38/00, 45/05, C07K 7/00, 14/82, C12N 15/00	A1	(43) International Publication Date: 31 December 1997 (31.12.97)			
,,,,,,,,		(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  Published  With international search report.  Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.			
(54) Title: ISOLATED, NUCLEIC ACID MOLECULES WHICH CODE FOR GAGE TUMOR REJECTION ANTIGEN, THE TUMOR					
REJECTION ANTIGEN, AND USES THEREOF					
(57) Abstract					
rejection antigen precursors are referred to as GAGE tumo	or reject liagnost	nucleic acid molecules which code for them, are disclosed. These tumor tion antigen precursors, and the nucleic acid molecules which code for it and therapeutic uses of the coding sequences and the tumor rejection tion antigens are also shown.			

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	Si	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ΑT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	T)	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	
BJ	Benin	ΙE	Ireland	MN	Mongolia	UA	Trinidad and Tobago Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	
BY	Belarus	IS	Iceland	MW	Malawi		Uganda
CA	Canada	IT	italy	MX	Mexico	US	United States of America
CF	Central African Republic	JP	Japan	NE	Niger	UZ	Uzbekistan
CG	Congo	KE	Kenya	NL.	Netherlands	VN	Vict Nam
СН	Switzerland	KG	Kyrgyzstan	NO		YU	Yugoslavia
CI	Côte d'Ivoire	KP	Democratic People's	NZ.	Norway	zw	Zimbabwe
СМ	Cameroon		Republic of Korea	PL	New Zealand Poland		
CN	China	KR	Republic of Korea	PT			
CU	Cuha	KZ	Kazakatan	RO	Portugal		
CZ	Czech Republic	LC	Saint Lucia	RU	Romania		
DE	Germany	Ц	Liechtenstein	SD	Russian Federation		
DK	Denmark	LK	Sri Lanka		Sudan		
EE	Estonia	LR	Liberia	SE	Sweden		
		1.454	Ciocria	SG	Singapore		

WO 97/49417

1

PCT/US97/10850

# ISOLATED, NUCLEIC ACID MOLECULES WHICH CODE FOR GAGE TUMOR REJECTION ANTIGEN, THE TUMOR REJECTION ANTIGEN, AND USES THEREOF

## RELATED APPLICATION

5 This application is a continuation-in-part of Serial No. 08/531,662, filed September 21, 1995, which is a continuation-in-part of copending Serial No. 08/370,648 filed January 10, 1995, which is a continuation-in-part of copending patent application Serial No. 08/250,162 filed on May 10 27, 1994, which is a continuation-in-part of Serial No. 08/096,039 filed July 22, 1993. Both of these applications are incorporated by reference.

## FIELD OF THE INVENTION

This invention relates to a nucleic acid molecule which

15 codes for a tumor rejection antigen precursor. More particularly, the invention concerns genes, whose tumor rejection antigen precursor is processed, inter alia, into at least one tumor rejection antigen that is presented by HLA-Cw6 molecules. The genes in question do not appear to be related to other known tumor rejection antigen precursor coding sequences. The invention also relates to peptides presented by the HLA-Cw6 molecules, and uses thereof. Also a part of the inventions are peptides presented by HLA-A29 molecules, and uses thereof.

## 25 BACKGROUND AND PRIOR ART

The process by which the mammalian immune system recognizes and reacts to foreign or alien materials is a complex one. An important facet of the system is the T lymphocyte, or "T cell" response. This response requires that T cells recognize and interact with complexes of cell surface molecules, referred to as human leukocyte antigens ("HLAs"), or major histocompatibility complexes ("MHCs"), and peptides. The peptides are derived from larger molecules which are processed by the cells which also present the HLA/MHC molecule. See in this regard Male et al., Advanced Immunology (J.P. Lipincott Company, 1987), especially chapters 6-10.

The interaction of T cells and HLA/peptide complexes is restricted, requiring a T cell specific for a particular combination of an HLA molecule and a peptide. If a specific T cell is not present, there is no T cell response even if 5 its partner complex is present. Similarly, there is no response if the specific complex is absent, but the T cell is This mechanism is involved in the immune system's response to foreign materials, in autoimmune pathologies, and in responses to cellular abnormalities. Much work has fo-10 cused on the mechanisms by which proteins are processed into the HLA binding peptides. See, in this regard, Barinaga, Science 257: 880 (1992); Fremont et al., Science 257: 919 (1992); Matsumura et al., Science 257: 927 (1992); Latron et al., Science 257: 964 (1992). Also see Engelhard, Ann. Rev. 15 Immunol. 12: 181-207 (1994).

The mechanism by which T cells recognize cellular abnormalities has also been implicated in cancer. For example, in PCT application PCT/US92/04354, filed May 22, 1992, published on November 26, 1992, and incorporated by reference, a family 20 of genes is disclosed, which are processed into peptides which, in turn, are expressed on cell surfaces, which can lead to lysis of the tumor cells by specific CTLs cytolytic T lymphocytes, or "CTLs" hereafter. The genes are said to code for "tumor rejection antigen precursors" or "TRAP" 25 molecules, and the peptides derived therefrom are referred to as "tumor rejection antigens" or "TRAs". See Traversari et al., Immunogenetics 35: 145 (1992); van der Bruggen et al., Science 254: 1643 (1991), for further information on this family of genes. Also, see U.S. patent application Serial 30 Number 807,043, filed December 12, 1991, now U.S. Patent No. 5,342,774.

In U.S. patent application Serial Number 938,334, now U.S. Patent No. 5,405,940, the disclosure of which is incorporated by reference, it is explained that the MAGE-1 gene codes for a tumor rejection antigen precursor which is processed to nonapeptides which are presented by the HLA-Al molecule. The reference teaches that given the known speci-

WO 97/49417

ficity of particular peptides for particular HLA molecules, one should expect a particular peptide to bind to one HLA molecule, but not to others. This is important, because different individuals possess different HLA phenotypes. As 5 a result, while identification of a particular peptide as being a partner for a specific HLA molecule has diagnostic and therapeutic ramifications, these are only relevant for individuals with that particular HLA phenotype. There is a need for further work in the area, because cellular abnormalities are not restricted to one particular HLA phenotype, and targeted therapy requires some knowledge of the phenotype of the abnormal cells at issue.

In U.S. Patent Application Serial Number 008,446, filed January 22, 1993 and incorporated by reference, the fact that 15 the MAGE-1 expression product is processed to a second TRA is disclosed. This second TRA is presented by HLA-C clone 10 molecules. The disclosure shows that a given TRAP can yield a plurality of TRAs.

U.S. Patent Application Serial Number 994,928, filed 20 December 22, 1992, and incorporated by reference herein teaches that tyrosinase, a molecule which is produced by some normal cells (e.g., melanocytes), is processed in tumor cells to yield peptides presented by HLA-A2 molecules.

In U.S. patent application Serial No. 08/032,978, filed 25 March 18, 1993, and incorporated by reference in its entirety, a second TRA, not derived from tyrosinase is taught to be presented by HLA-A2 molecules. The TRA is derived from a TRAP, but is coded for by a non-MAGE gene. This disclosure shows that a particular HLA molecule may present TRAs derived 30 from different sources.

In U.S. patent application Serial No. 08/079,110, filed June 17, 1993 and incorporated by reference herein, an unrelated tumor rejection antigen precursor, the so-called "BAGE" precursor is described. The BAGE precursor is not related to the MAGE family.

The work which is presented by the papers, patents, and patent applications cited <u>supra</u> deals, in large part, with

the MAGE family of genes, and the unrelated BAGE gene. It has now been found, however, that additional tumor rejection antigen precursors are expressed by cells. These tumor rejection antigen precursors are referred to as "GAGE" tumor 5 rejection antigen precursors. They do not show homology to either the MAGE family of genes or the BAGE gene. Thus the present invention relates to genes encoding such TRAPs, the tumor rejection antigen precursors themselves as well as applications of both.

Thus, another feature of the invention are peptides which are anywhere from 9 to 16 amino acids long, and comprise the sequence:

 $Xaa_{(1,2)}$  Trp Xaa Xaa Xaa Xaa Xaa Tyr (SEQ ID NO: 23)

15 where Xaa is any amino acid and Xaa<sub>(1,2)</sub> means that 1 or 2 amino acids may be N-terminal to the Trp residue. These peptides bind to, and/or are processed to peptide which bind to HLA-A29 molecules.

The invention is elaborated upon further in the disclo-20 sure which follows.

## BRIEF DESCRIPTION OF THE FIGURES

Figure 1 sets forth lysis studies using CTL clone 76/6.

Figure 2 shows tumor necrosis factor ("TNF") release assays obtained with various transfectants and controls.

25 Figure 3 compares lysis induced by cytolytic T lymphocytes of clone CTL 76/6. Peptides of varying length were tested, including SEQ ID NO: 4.

Figure 4 presents an alignment of the cDNAs of the six GAGE genes discussed herein. In the figure, identical regions are surrounded by boxes. Translation initiation sites and stop codons are also indicated. Primers, used in polymerase chain

5

reaction as described in the examples, are indicated by arrows.

Figure 5 sets forth the alignment of deduced amino acid sequences for the members of the GAGE family. Identical 5 regions are shown by boxes, and the antigenic peptide of SEQ ID NO: 4, is shown.

Figure 6 shows the results obtained when each of the GAGE cDNAs was transfected into COS cells, together with HLA-Cw6 cDNA. Twenty-four hours later, samples of CTL 76/6 were 10 added, and TNF release was measured after twenty-four hours.

Figure 7 compares the stimulation of CTL 22/23 by COS-7 cells, transfected with HLA-A29 cDNA, a MAGE, BAGE, or GAGE sequence, as shown. Control values are provided by MZ2-MEL.43 and COS cells, as stimulators.

15 Figure 8 presents results obtained from <sup>51</sup>Cr release studies, using various peptides including SEQ ID NO: 22 and various peptides derived therefrom.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

## Example 1

A melanoma cell line, MZ2-MEL was established from melanoma cells taken from patient MZ2, using standard methodologies. This cell line is described, e.g., in PCT Application PCT/US92/04354, filed May 22, 1992, published November 26, 1992, and incorporated by reference in its entirety.

25 Once the cell line was established, a sample thereof was

6

irradiated, so as to render it non-proliferative. These irradiated cells were then used to isolate cytolytic T cell clones ("CTLs") specific thereto.

A sample of peripheral blood mononuclear cells ("PBMCs")

5 was taken from patient MZ2, and contacted to the irradiated melanoma cells. The mixture was observed for lysis of the melanoma cells, which indicated that CTLs specific for a complex of peptide and HLA molecule presented by the melanoma cells were present in the sample.

- The lysis assay employed was a chromium release assay following Herin et al., Int. J. Cancer 39:390-396 (1987), the disclosure of which is incorporated by reference. The assay, however, is described herein. The target melanoma cells were grown in vitro, and then resuspended at 10' cells/ml in DMEM,
- 15 supplemented with 10 mM HEPES and 30% FCS, and incubated for 45 minutes at 37°C with 200  $\mu$ Ci/ml of Na( $^{51}$ Cr)O<sub>4</sub>. Labelled cells were washed three times with DMEM, supplemented with 10 mM Hepes. These were then resuspended in DMEM supplemented with 10 mM Hepes and 10% FCS, after which 100 ul aliquots
- containing  $10^3$  cells, were distributed into 96 well microplates. Samples of PBLs were added in  $100~\mu l$  of the same medium, and assays were carried out in duplicate. Plates were centrifuged for 4 minutes at 100g, and incubated for four hours at  $37^{\circ}C$  in an 8%  $CO_2$  atmosphere.
- Plates were centrifuged again, and 100 ul aliquots of supernatant were collected and counted. Percentage of 51Cr release was calculated as follows:

7

% 
$$^{51}$$
Cr release =  $(ER-SR)$  x 100  $(MR-SR)$ 

where ER is observed, experimental <sup>51</sup>Cr release, SR is spontaneous release measured by incubating 10<sup>3</sup> labeled cells in 200 ul of medium alone, and MR is maximum release, obtained by adding 100 ul 0.3% Triton X-100 to target cells.

Those mononuclear blood samples which showed high CTL activity were expanded and cloned via limiting dilution, and were screened again, using the same methodology. The CTL clone MZ2-CTL 76/6 was thus isolated. The clone is referred to as "76/6" hereafter.

The same method was used to test target K562 cells, as well as the melanoma cell line. Figure 1 shows that this CTL clone recognizes and lyses the melanoma cell line, i.e. MZ2-15 MEL but not K562. The clone was then tested against other melanoma cell lines and autologous EBV-transformed B cells in the same manner described <a href="mailto:supra">supra</a>. Figure 1 shows that autologous B cells, transformed by Epstein Barr Virus ("EBV") were not lysed, and that while MZ2-MEL 3.0 was lysed by CTL clone 76/6, the cell line MZ2-MEL.4F, a variant which does not express antigen F was not. Hence, the clone appears to be specific for this antigen.

The results presented <u>supra</u> are inconclusive as to which HLA molecule presents the TRA. The lysed cell line, i.e., 25 MZ2-MEL, is known to express HLA-A1, HLA-A29, HLA-B37, HLA-B44, HLA-Cw6, and HLA-C clone 10. In experiments not reported here but which follow the protocol of this example, a subline of MZ2-MEL was tested, which had lost expression of

8

HLA molecules A29, B44, and C clone 10. The subline was lysed, thus indicating that the presenting molecule should be one of A1, B37, or Cw6.

## Example 2

Further studies were carried out to determine if 76/6 also produced tumor necrosis factor ("TNF") when contacted with target cells. The method used was that described by Traversari et al., Immunogenetics 35: 145-152 (1992), the disclosure of which is incorporated by reference. Briefly, 10 samples of the CTL line were combined with samples of a target cell of interest in culture medium. After 24 hours, supernatant from the cultures was removed, and then tested on TNF-sensitive WEHI cells. Cell line MZ2-MEL.43, a subclone of the MZ2-MEL cell line discussed supra as well as in 15 the cited references, gave an extremely strong response, and was used in the following experiments.

## Example 3

The results from Example 2 indicated that MZ2.MEL.43 presented the target antigen of interest. As such, it was 20 used as a source of total mRNA to prepare a cDNA library.

Total RNA was isolated from the cell line. The mRNA was isolated using an oligo-dT binding kit, following well recognized techniques. Once the mRNA was secured, it was transcribed into cDNA, via reverse transcription, using an oligo dT primer containing a NotI site, followed by second strand synthesis. The cDNA was then ligated to a BstXI adaptor,

9

digested with NotI, size fractionated on a Sephacryl S-500 HR column, and then cloned, undirectionally, into the BstXI and Not I sites of pcDNA-I-Amp. The recombinant plasmid was then electroporated into DH5α <u>E</u>. <u>coli</u> bacteria. A total of 1500 pools of 100 recombinant bacteria were seeded in microwells. Each contained about 100 cDNAs, because nearly all bacteria contained an insert.

Each pool was amplified to saturation and plasmid DNA was extracted by alkaline lysis and potassium acetate preciption, without phenol extraction.

## Example 4

Following preparation of the library described in Example 3, the cDNA was transfected into eukaryotic cells. transfections, described herein, were carried out in dupli-15 cate. Samples of COS-7 cells were seeded, at 15,000 cells/ well into tissue culture flat bottom microwells, in Dulbecco's modified Eagles Medium ("DMEM") supplemented with 10% fetal calf serum. The cells were incubated overnight at 37°C, medium was removed and then replaced by 50  $\mu$ l/well of 20 DMEM medium containing 10% Nu serum, 400 μg/ml DEAE-dextran, and 100  $\mu$ M chloroquine, plus 100 ng of the plasmids. indicated supra, the lysis studies did not establish which HLA molecule presented the antigen. As a result, cDNA for each of the HLA molecules which could present the antigen 25 (A1, B37, Cw6) was used, separately, to cotransfect the cells. Specifically, one of 28 ng of the gene encoding HLA-Al, cloned into pCD-SRα was used, as were 50 ng of cDNA for

10

HLA-B37 in pcDNA-I-Amp, or 75 ng of cDNA for HLA-Cw6 in pcDNAI/Amp, using the same protocol as was used for transfection with the library.

Transfection was made in duplicate wells, but only 500 5 pools of the HLA-Cw6 transfectants could be tested in single wells. Following four hours of incubation at 37°C, the medium was removed, and replaced by 50  $\mu$ l of PBS containing 10% DMSO. This medium was removed after two minutes and replaced by 200  $\mu$ l of DMEM supplemented with 10% FCS.

10 Following this change in medium, COS cells were incubated for 24-48 hours at 37°C. Medium was then discarded, and 1000-3000 cells of CTL clone 76/6 were added, in 100 μl of Iscove's medium containing 10% pooled human serum supplemented with 20-30 U/ml of IL-2. Supernatant was removed 15 after 24 hours, and TNF content was determined in an assay on WEHI cells, as described by Traversari et al., Immunogenetics 35: 145-152 (1992), the disclosure of which is incorporated by reference.

The 1500 pools transfected with HLA-A1, and the 1500 20 pools transfected with HLA-B37 stimulated TNF release to a concentration of 15-20 pg/ml, or 2-6 pg/ml, respectively. Most of the HLA-Cw6 transfectants yielded 3-20 pg/ml, except for one pool, which yielded more than 60 pg/ml. This pool was selected for further work.

## 25 Example 5

The bacteria of the selected pool were cloned, and 600 clones were tested. Plasmid DNA was extracted therefrom,

11

transfected into a new sample of COS cells in the same manner as described supra, and the cells were again tested for stimulation of CTL clone 76/6. Ninety-four positive clones were found. One of these, referred to as cDNA clone 2D6, was 5 tested further. In a comparative test COS cells were transfected with cDNA clone 2D6 and the HLA-Cw6 cDNA, HLA-Cw6 cDNA alone, or cDNA 2D6 alone. Control cell lines MZ2-MEL F and MZ2-MEL F' were also used. TNF release into CTL supernatant was measured by testing it on WEHI cells, as referred to 10 supra. The number of surviving WEHI cells was measured by optical density after incubation of the cells with MTT. Figure 2 shows that the COS cells transfected with HLA-Cw6 and cDNA-2D6, and the cell line MZ2-MEL F' stimulated TNF release from CTL clone 76/6, indicating that HLA-Cw6 pre-15 sented the subject TRA.

## Example 6

The cDNA 2D6 was sequenced following art known techniques. A sequence search revealed that the plasmid insert showed no homology to known genes or proteins. SEQ ID NO: 1 20 presents cDNA nucleotide information for the identified gene, referred to hereafter as "GAGE". A putative open reading frame is located at bases 51-467 of the molecule. The first two bases of this sequence are from the vector carrying the cDNA sequence, and are thus not part of the cDNA itself.

12

## Example 7

Following sequencing of the cDNA, as per Example 6, experiments were carried out to determine if cells of normal tissues expressed the gene. To determine this, Northern 5 blotting was carried out on tissues and tumor cell lines, as indicated below. The blotting experiments used cDNA for the complete sequence of SEQ ID NO: 1. PCR was then used to confirm the results.

Table 1. Expression of gene GAGE

10	Normal tissues	
	PHA activated T cells CTL clone 82/30 Liver	- -
15	Muscle Lung Brain Kidney Placenta	<del>-</del> - - -
20	Heart Skin Testis	- - +
	Tumor cell lines	
25	Melanoma Lung Carcinoma Sarcoma Thyroid medullary carcinoma	7/16 1/6 0/1 0/1
	Tumor samples	
	Melanoma	1/1

13

## Example 8

Detailed analysis of normal tissues and tumors was carried out by applying polymerase chain reaction ("PCR") and the GAGE gene information described supra.

First, total RNA was taken from the particular sample, using art recognized techniques. This was used to prepare cDNA. The protocol used to make the cDNA involved combining 4 ul of reverse transcriptase buffer 5x, 1 ul of each dNTP, (10 mM), 2 ul of dithiothreitol (100 mM), 2 ul of dT-15 primer (20 um), 0.5 ul of RNasin (40 units/ul), and 1 ul of MoMLV reverse transcriptase (200 units/ul). Next, 6.5 ul of template RNA (1 ug/3.25 ul water, or 2 ug total template RNA) was added. The total volume of the mixture was 20 ul. This was mixed and incubated at 42°C for 60 minutes, after which 15 it was chilled on ice. A total of 80 ul of water was then added, to 100 ul total. This mixture was stored at -20°C until used in PCR.

To carry out PCR, the primers 5'-AGA CGC TAC GTA GAG CCT-3'

20 (sense)

and

5'-CCA TCA GGA CCA TCT TCA-3'
(antisense)

SEQ ID NOS: 2 and 3, respectively, were used. The reagents 25 included 30.5 ul water, 5 ul of PCR buffer 10x, 1 ul of each dNTP (10 uM), 2.5 ul of each primer (20 uM), and 0.5 ul of polymerizing enzyme "Dynazyme (2 units/ul). The total volume was 45 ul. A total of 5 ul of cDNA was added (this corre-

14

sponded to 100 ng total RNA). The mixture was combined, and layered with one drop of mineral oil. The mixture was transferred to a thermocycler block, preheated to 94°C, and amplification was carried out for 30 cycles, each cycle consisting 5 of the following:

first denaturation: 94°C, 4 min.
denaturation: 94°C, 1 min.
annealing: 55°C, 2 min.
extension: 72°C, 3 min.
final extension: 72°C, 15 min.

10

Following the cycling, 10 ul aliquots were run on a 1.5% agarose gel, stained with ethidium bromide.

cDNA amplified using the primers set forth supra yields a 238 base pair fragment. There is no amplification of contaminating genomic DNA, if present.

The results are presented in Table 2, which follows. They confirm that the only normal tissue which expresses GAGE is testis, whereas a number of tumors, including melanoma, lung, breast, larynx, pharynx, sarcoma, testicular seminoma, 20 bladder and colon express the gene. Thus, any one of these tumors can be assayed for by assaying for expression of the GAGE gene.

Table 2

15

# RT-PCR analysis of the expression of gene GAGE

# Heart Brain Liver Lung Kidney Spieen Lymphocytes Bone marrow Skin Naevus Melanocytes Fibroblasts Prostate Testis Ovary Breast Adrenals Muscle Placents Umbilical Cord

TUMORS	Cell lines	Tumor samples
Melanoma	40/63	45/146 (32%)
Lung cancer Epidermoid carcinoma Adenocarcinoma Small Cell Lung Cancer	6/23	10/41 (24%) 4/18 0/2
Breast cancer  Head and Neck tumor  Larynn  Pharynn		15/146 (10%) 6/15 (40%) 3/13
Sarcoma Testicular aeminoma	1/4	6/18 (33%)
Bladder cancer Prostate cancer		6/6 (100%) 5/37 (14%) 2/20
Colon carcinoma	5/13	0/38
Renal cancer Leukemia	0/6 3/6	0/45 0/19

16

## Example 9

The identification of the nucleic acid molecule referred to in the prior examples led to further work directed to the determination of tumor rejection antigens presented by HLA-5 Cw6 molecules, and derived from the GAGE gene.

The complete cDNA of GAGE in expression vector pcDNAI/
Amp was digested with restriction endonucleases NotI and
SpHI, and then with exonuclease III following supplier's
instruction (Erase-a-base System, Promega). This treatment
generated a series of progressive deletions, starting at the
3'end.

The deletion products were ligated back into pcDNAI/Amp, and then electroporated into E. coli strain DH5 $\alpha$ 'IQ, using well known techniques. The transformants were selected with 15 ampicillin (50 micrograms/ml).

Plasmid DNA was extracted from each recombinant clone and was then transfected into COS-7 cells, together with a vector which coded for HLA-Cw6. The protocols used follow the protocols described above.

The transfectants were then tested in the TNF release assay. This permitted separation of positive and negative clones. All the negative clones showed a deletion of the entire GAGE sequence. The smallest positive clone contained the first 170 nucleotides of SEQ ID NO: 1. The analysis of this sequence, supra, notes that the open reading frame starts at nucleotide 51. Thus, this fragment contains a sequence which encodes the first 40 amino acids of the GAGE TRAP.

17

## Example 10

Additional experiments were then carried out to define the region encoding the TRA peptide more precisely. Polymerase chain reaction ("PCR") amplification was used to do this.

Two primers were synthesized. The first primer was a 22-mer complementary to a sequence within the plasmid vector pcDNAI/Amp located upstream of a BamHI site. The second primer was a 29-mer containing at the 3'end nucleotides 102-119 of SEQ ID NO: 1, and at the 5'end an extension of 11 nucleotides containing an XbaI restriction site.

Following amplification, the PCR product was digested by BamHI and XbaI, and cloned into the BamHI-XbaI sites of plasmid pcDNA-3. The recombinant colonies were cotransfected into COS-7 cells with cDNA encoding HLA-Cw6, in accordance with Example 4, and a TNF release assay, also as described supra, was carried out, using CTL 76/6.

TNF release was observed, indicating that the "minigene" was processed to a TRA. The minigene, i.e., nucleotides 1-119 of SEQ ID NO: 1, the coding region of which runs from 20 nucleotides 51-119 encoded the first 23 amino acids of the cDNA of SEQ ID NO: 1. This information served as the basis for the next set of experiments.

## Example 11

Two peptides were synthesized, based upon the first 23 25 amino acids of SEQ ID NO: 1. These were:

Met Ser Trp Arg Gly Arg Ser Thr Tyr Arg Pro Arg Pro Arg Arg (SEQ ID NO: 2)

and

Thr Tyr Arg Pro Arg Pro Arg Arg Tyr Val Glu Pro Pro Glu Met Ile (SEQ ID NO: 3)

5 Each peptide was pulsed into COS-7 cells previously transfected with HLA-Cw6 cDNA, and combined with CTL 76/6 to determine if TNF release would be induced. Peptides (20 ug/ml) were added to COS-7 cells which had been transfected with the HLA-Cw6 cDNA twenty-four hours previously. After 10 incubation at 37°C for 90 minutes, medium was discarded, and 3000 CTLs were added in 100 microliters of medium, containing 25 units/ml of IL-2. Eighteen hours later, TNF content of supernatant was tested via determining toxicity on WEHI-164-The second peptide (SEQ ID NO: 3) was found to 13 cells. 15 induce more than 30 pg/ml of TNF, while the first peptide (SEQ ID NO: 2), was found to induce less than 10 pg/ml of The second peptide was used for further experiments. TNF.

## Example 12

Various peptides based upon SEQ ID NO: 3 were synthe20 sized, and tested, some of which are presented below. To
carry out these tests, <sup>51</sup>Cr labelled LB33-EBV cells, which are
HLA-Cw6 positive, were incubated with one of the following
peptides:

Tyr Arg Pro Arg Pro Arg Arg Tyr
25 (SEQ ID NO: 4)

Thr Tyr Arg Pro Arg Pro Arg Arg Tyr (SEQ ID NO: 5)

Tyr Arg Pro Arg Pro Arg Arg Tyr Val (SEQ ID NO: 6)

19

Thr Tyr Arg Pro Arg Pro Arg Arg Tyr Val (SEQ ID NO: 7)

Arg Pro Arg Pro Arg Arg Tyr Val Glu (SEQ ID NO: 8)

5 Met Ser Trp Arg Gly Arg Ser Thr Tyr Arg Pro Arg Pro Arg Arg (SEQ ID NO: 2)

The peptide concentration varied, as indicated in figure 3, and the ratio of CTL: LB33-EBV ("effector: target ratio"), was 10:1. <sup>51</sup>Cr release was determined after four hours of 10 incubation at 37°C. Levels of lysis for positive ("F'", MZ2-MEL.3.1), and negative ("F'"; MZ2-MEL.2.2.5) control cells are indicated, in figure 3.

In results not reported here, a second CTL was tested (CTL 82/31). This CTL was known to lyse cells presenting MZ2-F. It, too, lysed HLA-Cw6 positive cells following 25 pulsing with the peptide of SEQ ID NO: 4.

## Example 13

To find out whether the GAGE DNA set forth supra was unique, a cDNA library made with RNA from MZ2-MEL.43 (the same library that was used for the cloning of GAGE) was 5 hybridized with a probe derived from the GAGE cDNA. The probe was a PCR fragment of 308 base pairs between positions 20 and 328 of SEQ ID NO: 1. Twenty positive cDNAs were obtained. Six of them were entirely sequenced. They were all highly related to the GAGE sequence, but they were 10 slightly different from it. Two of the six clones were identical to each other, but all the others differed from Thus, five new sequences different from but each other. highly related to GAGE were identified. They are called GAGE-2, 3, 4, 5 and 6 (Figure 4). The fourteen other clones 15 were partially sequenced at the 5' end and their sequence corresponded to one of the six GAGE cDNAs.

The major difference between these cDNAs and GAGE-1 is the absence of a stretch of 143 bases located at position 379 to 521 of the GAGE sequence of SEQ ID NO: 1. The rest of the 20 sequences shows mismatches only at 19 different positions, with the exception of GAGE-3 whose 5'end is totally different from the other GAGE for the first 112 bases. This region of the GAGE-3 cDNA contains a long repeat and a hairpin structure.

The deduced GAGE-1 protein corresponding to a tumor rejection antigen precursor is about 20 amino acids longer than the 5 other proteins, whose last seven residues also differ from the homologous residues of GAGE-1 (Figure 5).

21

The rest of the protein sequences show only 10 mismatches. One of these is in the region corresponding to the antigenic peptide of SEQ ID NO: 4. The sequence of the peptide is modified in GAGE-3, 4, 5 and 6 so that position 2 is now W 5 instead of R.

## Example 14

To assess whether the change at position 2 affected the antigenicity of the peptide, cDNA of the 6 GAGE cDNAs were individually transfected into COS cells together with the 10 cDNA of HLA-Cw6, and the transfectants were tested for recognition by CTL 76/6 as described, supra. Only GAGE-1 and GAGE-2 transfected cells were recognized, showing that the modified peptide encoded by GAGE-3, 4, 5 and 6 was not antigenic in the context of this experiment. Sequence analysis of the 5' end of the 14 other clones mentioned supra, showed that 7 of them contained the sequence encoding the antigenic peptide, and thus probably corresponded to either GAGE-1 or GAGE-2.

## Example 15

The PCR primers used, <u>supra</u> to test the expression of GAGE in tumor samples do not discriminate between GAGE-1 or 2 and the four other GAGE cDNAs that do not encode antigen MZ2F. A new set of primers was prepared which specifically amplifies GAGE-1 and 2, and not GAGE-3, 4, 5 and 6. These 25 primers are:

22

VDE44 5'-GAC CAA GAC GCT ACG TAG-3' (SEQ ID NO: 9) VDE24 5'-CCA TCA GGA CCA TCT TCA-3' (SEQ ID NO: 10)

These primers were used as described, <u>supra</u>, in a RT-PCR reaction using a polymerase enzyme in the following tempera5 ture conditions:

4 min at 94°C
30 cycles with 1 min at 94°C
2 min at 56°C
3 min at 72°C

The results of this analysis are set forth in Table 3.

....

Table 3 Expression of GAGE genes by tumor samples and tumor cell lines

Histological type	Number of GAGE positive tumors			
	AJ GAGE genes"	GAGE-1 and 2*		
Tumor samples :	<del></del>			
Melanomas				
primary lesions	5/39	\$/39 (13M		
metastases	47/132	36/131 (27%		
Sarcomaa .	6/20	6/20 (30H		
Lung carcinomas NSCLG	14/65	12/64 (199		
Head and neck squamous call cardnomas	13/55	10/54 (199		
Prostatic carcinomas	2/20	2/20		
Hammary cardinomas	18/162	14/162 (9%		
Bladder carcinomas				
superficial	1/20	1/20		
infiltrating	5/26	\$/26		
Testicular seminomas	6/6	5/6		
Colorectal careinomas	0/43			
Leukemiss and lymphomas	0/25			
Renal cardnomas	0/46			
Tumor call fines				
Melanomás :	45/74	40/74 (54)		
Sarcomas	1/4	1/4		
Lung carcinomas ;				
sac.	7/24	7/24 (29)		
NSCLC	1/2	1/2		
Mesotheliomas	5/19	5/19 (26		
Head and neck squamous cell cardnoms	a 0/2	-		
Mammary cardnomas	1/4	0/4		
Bladder carcinomas	0/3			
Colon cardnomas ;	\$/13	\$/13		
Leukemias	3/6	1/6		
Lymphomas	0/6			
Renal carcinomas	0/6			

<sup>\*</sup> Expression of GAGE was tested by RT-PCR on total RNA with primers VDE-18 and VDE-24, detecting \$3 GAGE genes. No PCR product was observed whan these primers were essayed on DNA from MZ2-MEL.

\*\* Expression of GAGE-1 and 2 was tested by RT-PCR on total RNA with primers VDE-44 and VDE-24, which distinguish GAGE-1 and 2 from the four other GAGE genes. No PCR product was observed when these primers were assayed on DNA from MZ2-MEL.

In further work, new primers were designed which amplified all GAGE genes, to make sure that there was no expression of any of them in normal tissues. These primers are

VDE43 5'-GCG GCC CGA GCA GTT CA-3' (SEQ ID NO: 11)
5 VDE24 5'-CCA TCA GGA CCA TCT TCA-3 (SEQ ID NO: 10)

These were used exactly as for the PCR using the VDE44 and VDE24 primers. The results are shown in Table 4. They confirm that the normal tissues are negative, except for testis.

## Table 4

## Expression of GAGE genes in normal adult and fetal tissues

Adult tissues	GAGE expression <sup>e</sup>	
Adrenal pland	•	
Benign nazvus	•	
Bone marrow	•	
Brain	. •	
Breast	•	
Cerebellum ·	•	
Colon	•	
Heart	•	
Kidney	•	
Liver	•	
Lung	•	
Melanocytes	•	
Muscle	•	
Ovary	•	
Prostate	•	
Skin	•	
Splenocytes	•	
- Stomach	•	
Testis	<b>+</b>	
Thymocytes	•	
Urinal bladder	• .	
Uterus Piscenta	•	
PIRCENTE	•	
Umblical	•	
∞rd		
Fetal tissues*		
Fibroblasta	•	
Brain	•	
Liver	•	
Spleen	•	
Thymus	•	
Testis	<b>◆</b> '	

<sup>\*</sup>Expression of GAGE was tested by RT-PCR amplification on total RNA with primers VDE43 and VDE24 detecting all GAGE genes (Figure 7). Absence of PCR product is indicated by - and presence by +. No PCR product was observed when these primers were assayed on DNA from MZ2-MEL.

\*Fetal tissues derive from fetuses older than 20 weeks.

## Example 16

WO 97/49417

In work not reported here, it had been ascertained that cytolytic T cell clone CTL 22/23 (Van den Eynde, et al., Int. J. Cancer 44: 634-640 (1989), incorporated by reference) did 5 not recognize melanoma cell line MZ2-MEL.3.1. This melanoma cell line was reported by Van der Bruggen, et al., Eur. J. Immunol. 24: 2134-2140 (1994), to have lost expression of MHC molecules HLA-A29, HLA-B24, and HLA-cw·1601. Studies were undertaken to determine if transfection with one of these MHC 10 molecules could render the line sensitive to CTL 22/23. HLA-A29 was the first molecule tested. To do so, poly A' RNa was extracted from HLA-A29 cell line MZ2-MEL.43, using a commercially available extraction kit, and following the manufacturer's instructions. The mRNA was then converted to cDNA, 15 using standard methodologies, size fractionated, and then inserted unidirectionally, into the Bstx1 and NotI sites of plasmid pcDNA-I/Amp. The plasmids were electroporated into  $\underline{E}$ .  $\underline{coli}$  strain DH5 $\alpha'$ IQ, and selected with ampicillin (50 The bacteria were plated onto nitrocellulose filμg/ml). 20 ters, and duplicated. The filters were prepared, and hybridized overnight in 6xSSC/0.1% SDS/1x Denhardt's solution at 40°C, using 32P labelled probe:

## 5'ACTCCATGAGGTATTTC-3'

(SEQ ID NO: 19)

25 The probe is a sequence which surrounds the start codon of HLA sequences.

The filters were washed twice, at room temperature for 5 minutes each time in 6xSSC, and twice in 6xSSC at 43°C. Positive sequences were then screened with probe:

27

## 5'-TTTCACCACATCCGTGT-3'

5 (SEO ID NO: 20)

which had been labelled with <sup>32</sup>P. This sequence is specific for HLA-A29, as determined by reference to the Kabat Database of sequences and proteins of immunological interest, incorporated by reference. This database is available at the NCBI (USA), or on Web Sotle (Internet) WWW.NCBI.NLM.NIH.GOV. Teh filters were washed twice at room temperature for 5 minutes each time, at 6xSSC, followed by two washes, at 6xSSC (5 minutes per wash), at 42°C.

## Example 17

Once positive HLA-A29 clones were isolated, these were transfected into COS-7 using the DEAE-dextran chloroquine method set out <u>supra</u>. In brief, 1.5 x 10<sup>4</sup> COS-7 cells were treated with 50ng of plasmid pcDNA-I/Amp containing HLA-A29, and 100 ng of cDNA containing cDNA for one of the GAGE sequences mentioned <u>supra</u>, or one of the prior art MAGE or BAGE sequences in plasmid pcDNAα-I/Amp or pcDSRα-respectively. The transfectants were then incubated for 24 hours at 37°C.

The transfectants were then tested for their ability to stimulate TNF production by CTLs, using the assay explained 25 at the end of example 4, <a href="mailto:supra">supra</a>.

Figure 7, which presents the results of this drug, shows that high levels of TNF production were achieved using any of

GAGE-3, 4, 5 or 6 and HLA-A29 as transfectants. GAGE-1 and GAGE-2, in contrast, do not stimulate CTL clone 22/23, thus leading to the conclusion that GAGE 3, 4, 5 and 6 are processed to an antigen or antigens presented by HLA-A29 mole-5 cules and recognized by CTL 22/23.

## Example 18

The fact that GAGE-3, 4, 5 and 6 were processed to peptides presented by HLA-A29 cells, which GAGE-1 and GAGE-2 were not, suggested examination of the deduced amino acid sequences for those common to GAGE 3, 4, 5 and 6 and absent from GAGE-1 and GAGE-2.

The sequence:

Arg Ser Thr Tyr Tyr Trp Pro Arg Pro Arg Arg Tyr Val Gln (SEQ ID NO: 21)

15 was identified. The peptide was synthesized, lyophilized, and then dissolved in 1 volume DMSO, 9 volumes of 10 mM acetic acid in water. This methodology was used for the other peptides synthesized, discussed <a href="infra">infra</a>.

The peptide (SEQ ID NO: 21) was tested in a  $^{51}$ Cr release 20 experiment, following the method described  $\underline{\text{supra}}$ .

It was found that this peptide did provoke lysis. Successive deletions were prepared, and tested for their ability to provoke lysis, again using the 51Cr lytic assay. This work is depicted in Figure 8. It was found that the 25 shortest peptide to provoke lysis was

Tyr Tyr Trp Pro Arg Pro Arg Arg Tyr (SEQ ID NO: 22), which is common to all of GAGE-3 through 6.

Specifically, amino acids 10-18 of GAGE-3, and amino acids 9-17 of GAGE-4, 5 and 6 correspond to this peptide.

The members of the peptide family shown in Figure 8, and represented, e.g., by SEQ ID NOS: 21 and 22, do not accord 5 with the data presented by Toubert, et al., "HLA-A29 Peptide Binding Motif", Abstract No. 4183, Ninth International Congress of Immunology, July 23-29, 1995, San Francisco, CA, incorporated by reference. According to Toubert, et al., at the least a Phe residue is required at the third position of any peptide which binds to HLA-A29. As is shown herein, such is not the case.

## Example 19

A set of experiments were carried out to isolate and to clone genomic DNA sequences encoding GAGE TRAPS.

A library was made from genomic DNA isolated from the peripheral blood lymphocytes of patient MZ2. Isolation and preparation of the DNA was carried out in accordance with Wölfel et al., Immunogenetics 26: 178-187 (1987), incorporated by reference. The isolated DNA was then partially digested with the restriction enzyme Sau3A, and then fractionated using NaCl density gradient ultracentrifugation. This provides a fraction enriched in 10-20 kb fragments of DNA. See Grosveld et al., Nucl. Acids. Res. 10: 6715-6732 (1982). These fragments were dephosphorylated using alkaline phosphatase, and were then ligated into λ-Gem11 DNA, which had been digested with BamHI/EcoRI. Briefly, 2 ugs of the genomic DNA were mixed with 2 ugs of the λ phage DNA in a 10

PCT/US97/10850

ul volume, and incubated at 16°C overnight. 4 µl of the ligation mixture containing the ligated DNA was packaged, <u>in vitro</u>, in a commercially available phage packaging extract. The resulting phages were titrated on <u>E</u>. <u>coli</u> NM539 (a commercially available strain), in order to calculate the appropriate number of phages to plate out for screening. The resulting product was titrated onto cells of <u>E</u>. <u>coli</u> strain NM539.

## Example 20

- Approximately 33,333 recombinant phages were plated per plate, to give a total of 500,000 phages tested. A total of 20 μl of the packaging mixture was mixed with 1 ml of a suspension of E. coli NM539 in 10 mM MgSO<sub>4</sub>, to an OD<sub>600</sub> of 0.5. This mixture was then incubated, for 15 minutes at 37°C, and 15 then mixed with 15 ml of culture medium BTCYM containing 0.7% agarose at 45°C, and then plated onto agar plates containing BTCYM. The resulting mixture was incubated, at 37°C, overnight. The resulting phage plaques were used in hybridization experiments. Approximately 500,000 recombinant phage 20 plaques were immobilized on nylon membranes, and were then subjected to in situ hybridization, in accordance with Sambrook et al., Molecular Cloning: A Laboratory Manual (1989), incorporated by reference.
- The hybridization was carried out using a probe which 25 consisted of nucleotides 18 through 326 of SEQ ID NO: 1. The probe was prepared using the polymerase chain reaction and, as primers, a nucleotide sequence consisting of nucleotides

31

18-34 and the complement of nucleotides 309-326 of this sequence. The primers were used in a 30 cycle PCR run (1 cycle: 94°C for one minute, followed by 46°C for two minutes, then 72°C for three minutes), in a total volume of 100 ul, 5 which contained 10 ul of 10x concentrated Dynazyme buffer, 0.2 mM of each dNTP, 50 pmoles of each primer, and 2.5 units of Dynazyme DNA polymerase.

The probe was then purified via electrophoresis in low temperature melting agarose, as described by Sambrook et al., 10 supra. Following purification, the probe was radiolabelled with  $\alpha^{32}P$ , using a commercially available, random priming kit (radioactive nucleotide was  $\alpha^{32}P$  dCTP).

Once the probes were labelled, they were used in a hybridization buffer (10% sodium salt of dextran sulfate, MW 15 500,000; 1% SDS; 1M NaCl, and 50 ug/ml of denatured salmon sperm DNA). About 150 ng of <sup>32</sup>P labelled probe (approximately 1.6x10%cpm), were put into a total volume of 200 ml of this buffer. Approximately 500,000 immobilized plaques on filters were hybridized filter which was combined with the nylon 20 membrane containing at 65°C for about 15 hours. The membranes were then washed with 0.2xSSC, 0.1% SDS, at 65°C.

Following autoradiography, one positive clone was found. When excised, the insert was found to be about 11 kilobases long. Three fragments (175 base pairs, 4.5 kilobases, and 25 6.5 kilobases) resulted from treatment of the insert with the endonuclease SstI, and these were then subcloned into the plasmids pBluescript SK(-), and pTZ19R, both of which are commercially available. The fragments were sequenced in

32

their entirety, using commercially available enzymes, and primers 5'-labelled with  $[\gamma^{33}P]ATP$ . The sequence of the genomic clone is provided as SEQ ID NO: 24.

The foregoing examples show the isolation of nucleic 5 acid molecules which code for tumor rejection antigen precursors and tumor rejection antigens. These molecules, however, are not homologous with any of the previously disclosed MAGE and BAGE coding sequences described in the references set forth <a href="mailto:supra">supra</a>. Hence, one aspect of the invention is an isolated nucleic acid molecule which comprises the nucleotide sequence set forth in SEQ

ID NO: 1 as well as fragments thereof, such as nucleotides 1170, and 51-170, and any other fragment which is processed to
a tumor rejection antigen. The sequence of SEQ ID NO: 1 is
15 neither a MAGE nor a BAGE coding sequence, as will be seen by
comparing it to the sequence of any of these genes as described in the cited references. Also a part of the invention are those nucleic acid molecules which also code for a
non-MAGE and non-BAGE tumor rejection antigen precursor but
20 which hybridize to a nucleic acid molecule containing the
described nucleotide sequence, under stringent conditions.

The term "stringent conditions" as used herein refers to parameters with which the art is familiar. More specifically, stringent conditions, as used herein, refers to hybridization in 1M NaCl, 1% SDS, and 10% dextran sulfate.

This is followed by two washes of the filter at room temperature for 5 minutes, in 2xSSC, and one wash for 30 minutes in 2xSSC, 0.1% SDS. There are other conditions, reagents, and

so forth which can be used, which result in the same or higher degree of stringency. The skilled artisan will be familiar with such conditions, and, thus, they are not given here.

33

- It will also be seen from the examples that the invention embraces the use of the sequences in expression vectors, as well as to transform or transfect host cells and cell lines, be these prokaryotic (e.g., E. coli), or eukaryotic (e.q., CHO or COS cells). The expression vectors require 10 that the pertinent sequence, i.e., those described supra, be operably linked to a promoter. As it has been found that human leukocyte antigen HLA-Cw6 presents a tumor rejection antigen derived from these genes, the expression vector may also include a nucleic acid molecule coding for HLA-Cw6. In 15 a situation where the vector contains both coding sequences, it can be used to transfect a cell which does not normally express either one. The tumor rejection antigen precursor coding sequence may be used alone, when, e.g., the host cell already expresses HLA-Cw6. Of course, there is no limit on 20 the particular host cell which can be used. As the vectors which contain the two coding sequences may be used in HLA-Cw6 presenting cells if desired, and the gene for tumor rejection antigen precursor can be used in host cells which do not express HLA-Cw6.
- The invention also embraces so called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding

34

sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

To distinguish the nucleic acid molecules and the TRAPs

5 of the invention from the previously described MAGE and BAGE
materials, the invention shall be referred to as the GAGE
family of genes and TRAPs. Hence, whenever "GAGE" is used
herein, it refers to the tumor rejection antigen precursors
coded for by the previously described sequences. "GAGE

10 coding molecule" and similar terms, are used to describe the
nucleic acid molecules themselves.

The invention as described herein has a number of uses, some of which are described herein. First, the invention permits the artisan to diagnose a disorder such as melanoma, 15 characterized by expression of the TRAP, or presentation of the tumor rejection antigen. These methods involve determining expression of the TRAP gene, and/or TRAs derived therefrom, such as a TRA presented by HLA-Cw6. In the former situation, such determinations can be carried out via any 20 standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labelled hybrid-In the latter situation, assaying with ization probes. binding partners for complexes of TRA and HLA, such as antibodies, is especially preferred. An alternate method for 25 determination is a TNF release assay, of the type described supra. To carry out the assay, it is preferred to make sure that testis cells are not present, as these normally express GAGE. This is not essential, however, as one can routinely

35

differentiate between testis and other cell types. Also, it is practically impossible to have testis cells present in non-testicular sample.

The isolation of the TRAP gene also makes it possible to isolate the TRAP molecule itself, especially TRAP molecules containing the amino acid sequence coded for by SEQ ID NOS: 2-6. These isolated molecules when presented as the TRA, or as complexes of TRA and HLA, such as HLA-Cw6 or HLA-A29 may be combined with materials such as adjuvants to produce vaccines useful in treating disorders characterized by expression of the TRAP molecule.

Exemplary adjuvants include Freund's complete and incomplete adjuvant, killed B. pertussis organism, "BCG", or Bacille Calmente-Guerin, Al (OH), muramyl dipeptide and its derivatives which may be emulsified in metabolizable oils, such as squalene, monophosphoryl lipid A (MPL), keyhold limpet hemocyanin (KLH), saponin extracts such as QA-7, QA-19, and QA-21 (also referred to as QS-21), these having been described in U.S. Patent No. 5,057,540 to Kensil, et al., incorporated by reference, MTP-MF59, N-[1-(2,3-dioleoyloxy) propyl]-N,N,N-trimethylammonium methyl sulfate (DOTAP), the cationic amphiphile DOTMA, the neutral phospholipids such as DOPE, and combinations of these. This listing is by no means comprehensive, and the artisan of ordinary skill will be able to augment this listing. All additional adjuvants are encompassed herein.

In addition, vaccines can be prepared from cells which present the TRA/HLA complexes on their surface, such as non-

WO 97/49417

PCT/US97/10850

36

proliferative cancer cells, non-proliferative transfectants, etcetera. In all cases where cells are used as a vaccine, these can be cells transfected with coding sequences for one or both of the components necessary to provide a CTL response, or be cells which express both molecules without transfection. Further, the TRAP molecule, its associated TRAs, as well as complexes of TRA and HLA, may be used to produce antibodies, using standard techniques well known to the art.

- When "disorder" is used herein, it refers to any pathological condition where the tumor rejection antigen precursor is expressed. An example of such a disorder is cancer, melanoma in particular. Melanoma is well known as a cancer of pigment producing cells.
- As indicated, <u>supra</u>, tumor rejection antigens, such as the one presented in SEQ ID NO: 4 are also a part of the invention. Also a part of the invention are polypeptides, such as molecules containing from 8 to 16 amino acids, where the polypeptides contain the amino acid sequence set forth in
- 20 SEQ ID NO: 4. As the examples indicate, those peptides which are longer than the octamer of SEQ ID NO: 4 are processed into the tumor rejection antigen of SEQ ID NO: 4 by the HLA-Cw6 presenting cancer cells, and presented thereby. The presentation leads to lysis by cytolytic T lymphocytes pres-
- 25 ent in a body fluid sample contacted to the cells presenting the complex. Similarly, the peptides longer than SEQ ID NO: 22, such as SEQ ID NO: 21, are processed to the appropriate TRA, and are presented by cancer cells, such as HLA-A29

37

positive cells.

Thus, another feature of the invention are peptides which are anywhere from 9 to 16 amino acids long, and comprise the sequence:

5 Xaa Xaa Trp Xaa Xaa Xaa Xaa Trp
(SEO ID NO: 23)

where Xaa is any amino acid. These peptides bend to, and/or are processed to peptides which bind to HLA-A29 molecules. The fact that these peptides are processed to the tumor 10 rejection antigen, is indicated by the examples.

This property may be exploited in the context of other parameters in confirming diagnosis of pathological conditions, such as cancer, melanoma in particular. For example, the investigator may study antigens shed into blood or urine, observe physiological changes, and then confirm a diagnosis of melanoma using the CTL proliferation methodologies described herein.

On their own, peptides in accordance with the invention may be used to carry out HLA-typing assays. It is well known that when a skin graft, organ transplant, etc., is necessary one must perform HLA typing so as to minimize the possibility of graft rejection. The peptides of the invention may be used to determine whether or not an individual is HLA-Cw6 positive, so that appropriate donors may be selected. This type of assay is simple to carry out. The peptides of the invention are contacted to a sample of interest, and binding to cells in that sample indicates whether or not the individual from which the sample is taken is HLA-Cw6 positive. One

38

may label the peptides themselves, conjugate or otherwise bind them to linkers which are labeled, immobilize them to solid phases, and so forth, so as to optimize such an assay. Other standard methodologies will be clear to the skilled 5 artisan, and need not be presented herein.

Therapeutic approaches based upon the disclosure are premised on a response by a subject's immune system, leading to lysis of TRA presenting cells, such as HLA-Cw6 cells. One such approach is the administration of CTLs specific to the complex to a subject with abnormal cells of the phenotype at issue. It is within the skill of the artisan to develop such CTLs in vitro. Specifically, a sample of cells, such as blood cells, are contacted to a cell presenting the complex and capable of provoking a specific CTL to proliferate. The target cell can be a transfectant, such as a COS cell of the type described supra. These transfectants present the desired complex on their surface and, when combined with a CTL of interest, stimulate its proliferation. COS cells, such as those used herein are widely available, as are other suitable host cells.

To detail the therapeutic methodology, referred to as adoptive transfer (Greenberg, J. Immunol. 136(5): 1917 (1986); Riddel et al., Science 257: 238 (7-10-92); Lynch et al., Eur. J. Immunol. 21: 1403-1410 (1991); Kast et al., Cell 25 59: 603-614 (11-17-89)), cells presenting the desired complex are combined with CTLs leading to proliferation of the CTLs specific thereto. The proliferated CTLs are then administered to a subject with a cellular abnormality which is

characterized by certain of the abnormal cells presenting the particular complex, where the complex contains the pertinent HLA molecule. The CTLs then lyse the abnormal cells, thereby achieving the desired therapeutic goal.

The foregoing therapy assumes that at least some of the subject's abnormal cells present the relevant HLA/TRA complex. This can be determined very easily, as the art is very familiar with methods for identifying cells which present a particular HLA molecule, as well as how to identify cells expressing RNA of the pertinent sequences, in this case a GAGE sequence. Once cells presenting the relevant complex are identified via the foregoing screening methodology, they can be combined with a sample from a patient, where the sample contains CTLs. If the complex presenting cells are lysed by the mixed CTL sample, then it can be assumed that a GAGE derived, tumor rejection antigen is being presented, and the subject is an appropriate candidate for the therapeutic approaches set forth supra.

Adoptive transfer is not the only form of therapy that is available in accordance with the invention. CTLs can also be provoked in vivo, using a number of approaches. One approach, i.e., the use of non-proliferative cells expressing the complex, has been elaborated upon supra. The cells used in this approach may be those that normally express the complex, such as irradiated melanoma cells or cells transfected with one or both of the genes necessary for presentation of the complex. Chen et al., Proc. Natl. Acad. Sci. USA 88: 110-114 (January, 1991) exemplifies this approach, show-

40

ing the use of transfected cells expressing HPV E7 peptides in a therapeutic regime. Various cell types may be used. Similarly, vectors carrying one or both of the genes of interest may be used. Viral or bacterial vectors are espesially preferred. In these systems, the gene of interest is carried by, e.g., a Vaccinia virus or the bacteria BCG, and the materials de facto "infect" host cells. The cells which result present the complex of interest, and are recognized by autologous CTLs, which then proliferate. A similar effect can be achieved by combining the tumor rejection antigen or the precursor itself with an adjuvant to facilitate incorporation into HLA-Cw6 presenting cells which then present the HLA/peptide complex of interest. The TRAP is processed to yield the peptide partner of the HLA molecule while the TRA is presented without the need for further processing.

Other aspects of the invention will be clear to the skilled artisan and need not be repeated here.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there 20 is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, it being recognized that various modifications are possible within the scope of the invention.

504 10 NO. 29

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	<del></del>
GAGCTOGCTG	CAGCCITGAC	CTCCTGGGCT	CAAGOGCICC	TOOCACCTCA	50
GCCTCCTGAG	TAGCTGTGAG	TATAGGIACA	TGCCACCATG	CNCAGCTAAT	100
TTTTCGATGG	TITTITIGIT	TGITTTTTGT	AGTGATGAGA	TTTTCTCATG	150
TTGCTTAGGC	TGGTCTCGAA	GICCIGAGCI	CAGGTGATCT	GGCCAGCTCA	200
GCCTCCCAAA	ATACTAGGAT	TACAGGOGIG	ANITGGCCTG	GICTGGITTT	250
TCTTATATAG	GGGICTTATC	TATATAAAGA	CTAAAGTTAA	TCTGTGCCTT	300
TGTGCGGGTG	GGCTAAGAGC	ATGATGACIT	TTATCATTCT	ATTGATTTAA	350
AGAAAACIGI	CCTTGACTTA	CCAGTGTGTA	AGTOCATGAA	AGCATAATTC	400
TGTTGAAAGC	ATATATIGIT	AATGGGIGIT	GGGAACCGTG	CACTTTCCCC	450
TGCTGTGGGA	CCATCTCCTT	GGAGGTACCT	TICATCIGIT	TICICAACIC	500
CAAACATCTT	AGGACCATGG	GITGIGACIG	GTAGGACTAT	GIATCTTCCT	550
GCTTTCAAGA	CCCACTATAT	TTTCACGTCG	TGTCACTCTG	CCICICCICI	600
TTCCCTAATA	CIGICACITC	ACCCTCTGCG	ATTCTGATGC	TACAAATGAT	650
AGATATOGTT	TTAGCATTTT	CITACGGGIC	CTAGCGATTC	TATTCATTTT	700
TCTTTCAGTC	TCTTTCTCTG	ACTIGITICAC	ATTGAACAAT	TTCCTTTTCG	750
GATAGGITGC	TATTICIGIT	TICGCAGGIG	GITTACCIGI	CTTCCCAGCC	800
AGTCACAGTG	GICCLICICC	CCATGGTGGG	TOOGGGGCAA	GAGAGGGCCC	850
TGGGTTGGGG	GIGGGGITCA	GITGAAGATG	GGGTGAGTTT	TGAGGGGAGC	900
ACTACTIGAG	TOOCAGAGGC	ATAGGAAACA	GCAGAGGGAG	GIGGGATTCC	950
CITATOCICA	ATGAGGATGG	GCATGGAGGG	TITIGGGGGGT	GCCCTCGGA	1000
ACGGCAGCCC	TOOCCAGOOC	ACAGCCGCGC	ATGCTCCCTG	NICCOGCCIC	1050
AGIGOGCAIG	TICACIGGGC	GICTICIGOC	COCCCCTTC	GCCCACGIGA	1100
AGAACGCCAG	GGAGCIGIGA	GGCAGTGCTG	TETEGTTCCT	GCCCTCCCGA	1150
CTCTTTTTCC	TCTACTGAGA	TTCATCIGGT	AGGIGIGCAG	GCCAGTCATC	1200
••••			AGGGCCTCGG		1250
			GGAGAAGGGC		1300
			GCCACCGGCG		1350
			GGCCGTGGAG		1400
•			TGTTGGGGGG		1450
			CAGATICCCT	<del>-</del>	1500
			GCCIGGCACC		<sub>.</sub> 1550
			TOGAGTGCCG		1600
TGAGAAGCAC	TOCAAGGTCT	CACCTCCCCC	ATGGAAGGTC	CGAAAACAGT	1650
			ACCAAACITG		1700
			CCCAAAGCAG	· <del>-</del>	1750
_			GICCACTIGI		1800
CICATCAGAA	ATGACCICIG	TCTGCGGGGC	GCAGTGGCGC	TOGOCTAOGT	1850

42
3.3 updated LK new Sequence

10 20 30 40 50 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1900 10000000000000000000000000000000						
AGTICOCAGIT         ACTIGAGGACA         CITCAGGIGG         AGGATICCTT         GAGGGGAG         1900           TOCAGGGTICA         AGTIGAGCIGT         GATTAGCOCG         CITCACTOCA         GCCTGAGCAA         1950           CACAGGGATA         CCCGGIGTOC         AAAAGAAATT         TAGAAAAAAA         TGTCCTCTGC         2000           CTITITGCCAC         ACGCTTAAG         ATCATTGCTT         TGCCAGCTGG         GCCAGCAGAA         2050           GIGGCTTTIGT         AGGGATATTT         CAAAAGTAAA         ACGCGACAG         CTCCTATGCG         2100           ACTTATTTTG         AGAGTATTATA         GCAGCCAGCAG         CTCCAATCA         2200           CGGTTGCTG         TAGACCAAGG         CCCTATGTAC         ACGTCTCTA         AAGATTGGC         2250           CGTTATTGCCG         TGAGCCAAGG         CCCTATGTAC         AGGTTTTCTAC         AATTAATTTGCC         TGATTCTGCA         AAGGAAACAT         2450           CGTCTTTCCT         GCTCATAAAA         AATGAGCATG         GCATCTCTAC         AAGCAAACAT         2450           ATGACTTCTT         TCTCCACACA         GCATCTTATT         TTTTTCCACACAC         GCATCTTATT         TTTTTTCCACACACA         GCATCTACATT         TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	10	20	30	40	50	
AGTICOCAGIT         ACTIGAGGACA         CITCAGGIGG         AGGATICCTT         GAGGGGAG         1900           TOCAGGGTICA         AGTIGAGCIGT         GATTAGCOCG         CITCACTOCA         GCCTGAGCAA         1950           CACAGGGATA         CCCGGIGTOC         AAAAGAAATT         TAGAAAAAAA         TGTCCTCTGC         2000           CTITITGCCAC         ACGCTTAAG         ATCATTGCTT         TGCCAGCTGG         GCCAGCAGAA         2050           GIGGCTTTIGT         AGGGATATTT         CAAAAGTAAA         ACGCGACAG         CTCCTATGCG         2100           ACTTATTTTG         AGAGTATTATA         GCAGCCAGCAG         CTCCAATCA         2200           CGGTTGCTG         TAGACCAAGG         CCCTATGTAC         ACGTCTCTA         AAGATTGGC         2250           CGTTATTGCCG         TGAGCCAAGG         CCCTATGTAC         AGGTTTTCTAC         AATTAATTTGCC         TGATTCTGCA         AAGGAAACAT         2450           CGTCTTTCCT         GCTCATAAAA         AATGAGCATG         GCATCTCTAC         AAGCAAACAT         2450           ATGACTTCTT         TCTCCACACA         GCATCTTATT         TTTTTCCACACAC         GCATCTTATT         TTTTTTCCACACACA         GCATCTACATT         TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1234567890	1234567890	1234567B90	1234567890	1234567890	
TCGAGGCTCC         AGTGAGCTGT         GATCAGCGCC         CTGCAGCTCA         1950           CACAGGGATA         CCGGGTGTCC         AAAAGAAATT         TTAGAAAAAAA         TGTCCTCTGC         2000           CACAGGGATA         CCGGTTGTC         AAAAGAAATT         TTAGACACTGC         2000         2000           GTGGTTTGT         AGGCACTCAG         ACGCTACATTAG         CCAGTTAGCT         2000         2000           ACTTATTTTC         AGGCACACTGT         GCAGCACAA         2050         2000         2000           CATGGAAGT         ATCGAATATA         ACAGCTAGCA         ACGCTACACAC         2250         2250           CATTATTGCC         TAGACCACAC         CCCTATGTAC         ACCTCTGA         AATGATTGCC         2250           ATTATTGCC         TAGACCACAC         CCCTATGTAC         ACCTCTGAAA         AATGATTGCC         2250           AATTAATTTT         TTGGATTACC         TCGTTCATT         AATGATTCTT         TATTATTGCCC         2300           CCTATGCGG         TGGATTACACA         AACGTTATATA         ACCTCTATACACACA         AACGTTACACACAACAACAACAACAACAACAACAACAACAACAA						1900
CACAGOGATA         COCCITICO         AAAAGAAATT         TAGAAAAAAA         TIGICCICIEC         2000           CITTIGOCAC         ACGOCITAAG         ATGATICCIT         TOCCAGOCIG         GOCAGCAGAA         2050           GIGGOTTIGT         AGCACICAG         ACACGITACAT         CACGITATICT         TAACICIGGG         2100           ACTITATITIG         AGGACATCIAG         ACAGOCTIGI         GGAGOCACG         TICCCAATCA         2200           CGGITGICIG         TITTCAGIGT         GAAATATGAG         TICGCAGCAA         AACTICATCACC         2250           ATTATTGCC         TAGACCAAGG         CCCTATGTAC         AACCTCCTCA         AATGATTGCC         2300           CCTATGCGGG         TGAGICATTA         AACGITAATT         CGATGITTC         TATTAGIGAA         2350           AATTAATTT         TGICATACAG         TCCTTACAAA         AATGAGCAT         AGCATCACAT         AATGACCAAT         2450           ACCTACTACAC         GAGTATAATA         ACCCCCAAAG         TCCTCGTCA         ACCACCATAA         2450           TTCCAACAC         GAGTATAATA         GCCCCAAAG         TCCTCGTCC         TCACCTITICC         2500           CACAGTAACC         TAGAGCAAA         ATGACCATAC         TTATTTTTTAT         2700         TTATTTTTTTTTAT         2750						
CTITITICCAC ACCCITAAG ATGATTGCTC TECCAGCCTG GCCACCAGAA 2050 GTGGCTTTGT AGGCACTCAG ACACCGTACA CACGTATGCT TAACTCTGGG 2100 ACTTATTTTG AGAGTATTT CAAAAGTAAA ACGCCAAGTT AACATTTATC 2150 ACTTATTTTG AGAGTATTT CAAAAGTAAA ACGCCAAGTT AACATTTATC 2150 ACTTATTTTG AGAGTATTT CAAAAGTAAA ACGCCAAGTT AACATTTATC 2200 CCGGTTGCTG TTTTCAGTGT GAAATATCAG TTGCCGAGCA AGATCAACCT 2250 ACTTATCGCC TAGACCAAGG CGCTATGTAC ACCCTCTGA AATGATTCAG 2300 CCTATGCGGG TGAGTGCTTA AACGTTAATT CGAAGTTTTC TATTAGTACA 2350 ACTTACTCGG TGAGTGCTTA AACGTTAATT CGAAGTTTTC TATTAGTACA 2350 ACTTACTCGG TGAGTGCTTT AACGTTCATT AGTGTGGAAA TGCTCATAAAA 2400 GGTCTTTCCT GCTCATAAAA AATGACGATG GCATCTCAG AAGGAAACAT 2450 GGTCTTTCCT GCTCATAAAA AATGACGATG GCATCTCAG AACGAAACAT 2450 ATGACTTCTT TCTCCGCCTT TGTTTGTTAA TGACCAGTTG TACCACATGTA 2550 ATGACTCTTT TCTCCGCCTT TGTTTGTTAA TGACCAGATTG TACCACATGTA 2550 ACTTGCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTTCT 2600 CCACAGTAACC TCCCTGTGGG TGGAGTAACC TTATTGCGCA TACGACCATAG 2650 ACTTGGACAAA ATGTCTTTAG GCCTATAGTAA TACCAAAACT TATTTTTTTT 2770 TCCACACCC ACACGTATTC CCCACCCCGA GCAGTTCAGT TACACATGTAT 2770 TCTGGTGTATA TATGTAAAAT TTTGTATCAA TAACCAAAACT TATTTTTTTT 2750 TTTCCACACCC ACACGTATTC CCCACCCCGA GCAGTTCAGG TCACGTTTTTTT 2750 AACCACCAAC ACCGTAACAA GGGGAACCAG CCACTCCAG GTCAAGGTGC 2800 AACCACCAAC ACCGTGAACA GGGGAACCAG CCACTCCAG GTCAAGGTGC 2800 AACCACCAAC ACCGTGACAA GGGGAACCAG CCACTCCCAG GTCAAGGTGC 2900 CCCGCAAAGCGA ACACAACGT CTCCTGGTGT GTCGTGTTCCT GTGTGTTTCTCT 2950 ACCACCACAAA AAAAACAATCG AAACAATCG TCCCTGTGTG GTCGTGTTCTCT GTGTGTTTCTCT 2950 ACCACCACAC ACCCTTACCT GTGTGTTTCAC TTAGCAACCA ACACAACAC CCCTCCACACAC TCACACAAAACC TCCACACAAAACC AACACAACAACA AACACATTACA ACCTTTACCAC TCCACACAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACAAAAACC TCCACACACA						
GRECTITIGT AGGACTICAG ACACCGTACA CACGTATGCT TAACTCTGGG 2100 ACTTATITITG ACAGTATITT CAAAAGTAAA ACGCCAGGTT AACATTTATC 2150 CATGGAAGTG ATCGAATATA GCACCCTGT GGAGCGCACG TITCCCAATCA 2200 CCGTTGTCTG TITTCAGGTG GAAATATGAG TITGCCGAGGA ACATCGACCT 2250 ATTATTGCCC TAGACCAAGG CGCTATGTAC ACCCTCTGA AATGATTGGG 2300 CCTATGCGGG TGAGTCCTTA AACGTTAATT CGATGTTTC TATTAGGAG 2300 CCTATGCGGG TGAGTCCTTA AACGTTAATT CGATGTTTTC TATTAGGAG 2300 AATTAATTTT TGTGATAGCG TCGTTCCATT AGTGGGAAA TCCTCGATAAA 2400 GGICTTTCCT CCTCATAAAA AATGAGGATG GCATCTCATG AACGAAACAT 2450 TGATTCTGGA GGATTTTTT TTTTCCTCTC GIGTTCTTCA GCTTTTCCCC 2500 ATGACTTCTT TCTCCGCCTT TGTTTGTTAA TGACACATGT TACACATGTA 2550 ATGACTACTT TCTCCGCCTT TGTTTGTTAA TGACACATGT TACACATGTA 2550 ATGACTACACA GAGTATATA CCCCCCAAAA TCCTCGTGCG TCACTTTTCT 2600 CCACAGTAACA TCCTCGTGGG TGGAGTAACA TCCTCGTGCG TCACTTTTCT 2600 CCACAGTAACA TACTCTTAG GCTTAGTTAG TACACAATAG 2550 AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG TACACAATAG 2600 CCACAGTAACA ACCTGAAGAA TTTTGTACAA TAACCAAACT TATTTTTTTT 2750 TTCCACACCC ACACGTATC CCCACCCCA GCACTTCACG TCACTTTTCT 2750 TTCCACACCC ACACGTATC CCCACCCCA GCACTTCACG TCACTTTTTTT 2750 TTCCACACCC ACACGTATC CCCACCCCAA TAACCAAACT TATTTTTTTTTT						
ACTTATTITIG ACAGIATITI CAAAAGTAAA ACGCCAAGTT AACAITTATC CATGGAAGG ATCGAATATA GCAGCCTGT GGAGCGACG TITCCCAATCA CCGTTGTCTG TITTCAGTGT GAAATATGAG TITGCCGAGCA ACATCGACCT ATTATTGCCC TAGACCAAGG CCCTATGTAC AGCCTCCTCA AATGATTGCG CCTATGCCCG TGAGTCCTTA AACGTTAATT CGATGTTTCT TATTAGTAGA CATTAATTTT TGTGATAGCG CCCTATAGTAC AGCCTCCTCA AATGATTGCG CCTATCCCCG TGAGTCCTTA AACGTTAATT CGATGTTTCT TATTAGTAGA 2350 AATTAATTTT TGTGATAGCG TCGTTCATT AGTGTGGAAA TCCTCATAAA CGTCTTTCCT CCTCATAAAA AATGACGATG CCATCTCATG AACGAAACAT CGATTCTTGA GCATTTTTT TTTTCCTCTC GTGTTCTCA GCTTTTCCCC 2500 ATGACTTCTT TCTCCCCCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2450 TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTCG TCACTTTTCT 2600 CACAGIAACC TCCCTGTGGG TGAGTAACA TCCTCGTGCG TCACTTTTCT 2600 CACAGIAACC TCCCTGTGGG TGGAGTAACC TTATTTGTACA TAACCAAACT TATTTTTTTT 2750 ACTTGGACAAA ATGTCTTTAG GCTTAGTTAG GACCACAAAT ACCTATGTAT 2700 CCCACACAAA ATGTCTTTAG GCTTAGTTAG GACCACAAAT ACCTATGTAT 2750 TCTGTGTATA TATGTAAAAT TTTTGTACAA TAACCAAACT TATTTTTTAT 2750 TCTGTGTATA TATGTAAAAA TTTTGTAACAA TAACCAAACT TATTTTTTAT 2750 AACCACCACAC ACCTGAACAA GGGGAACCAG CAACTCAACG TCACGATCCT 2800 CCCAAAGCAAC ACCTGAACAA GGGGAACCAG CAACTCAACG TCACGATCCT 2800 CCCAAAGCACA ACAACAACCT CTCCTGGTGT GTGCGTGTGT GTGTGTTCTGT 2950 CGCAAAAGCAA ACAACAACCT CTCCTGGTGT GTGCGTGTGT GTGTGTTCTGT 2950 CGCGAAAGCAA AAAACAATGC GTGTGTTAGC CATTTGCACA TACCACACAAG 3000 CCCGAAAGCAAA AAAACAATCC GTGTGTTAGC CATTTGCACA TACCACACAAG 3000 CCCGAAAACCA CACTTCACCT GTGTGTTTCACA TACCAACACC 3050 ACCAGCCTAT GTAGTTTGCA GCTTACCTTA GCCAAAATCC TCACTATCAT 3100 AAAAGTTCTC CACTTTATCA GTGTGTTCCT GGAAAATCC TCACTATCAT 3250 CGCTGAAAATG CTTAAAACT CACTTTACAC TCACAACC AACTTCACAC TCCACACACA						_
CATGAAGIG ATCGAATATA GCAGCCCIGT GGAGCGACG TITCCCAATCA CGGITGICIG TITTCAGIGT GAAATATGAG TIGCCGAGGA AGATCGACCT 2250 ATTATTGGCC TAGACCAAGG CCCTATGTAC AGCCTCCTGA AATGATTGGG AACTTAATTT TITGATAGCG TCGITGCATT AGCGTGATTAC AGCCTCCTGA AATGATTGGG AACTTAATTTT TITGATAGCG TCGITGCATT AGCGTGATAA ACCTTAATTT CGATGTTCAA AACGAAACAT CGGICITTCCT GCTCATAAAA AATGAGGATG GCACTCCATG AACGAAACAT TGATTCTGGA GCATTTTTT TITTCCCTCTC GIGITCTTCA GCTTTTGCCC ATGACCTTCTT TCTCCGGCTT TGTTTGTTAA TGACGAGATTG TACACATGTAA ATGACTTCTT TCTCCGGCTT TGTTTGTTAA TGACGAGATTG TACACATGTA 2550 ACCAGGAAACC TCCCTGTGGG TGCAGTTAACC TCCCTGTGCG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGGG TGCAGTAACA TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACA ATGTCTTTAG GCTTAGTATA TATACCAAAAT ACCTATGTAT 2770 TCTGGTGATAT TATGTAAAAAT TTTGTATCAA TAACCAAAACT TATTTTTTAT 27750 TCTGGTGATAT TATGTAAAAAT TTTGTATCAA TAACCAAAACT TATTTTTTAT 27750 TTTGCACACCA ACCTGAAGAA GGGGAACCAG CAACTCAACG TCACGATCCT 2850 GCCAAAAGCGA ACCAGCACAG GGATGACGGA GCACTTCACG TCAACGTCA 2890 GCGAAAAGCGA ACAACAACGT CTCCCTGGTGT GTGCGTGTGT GTGTGTTTCTT GCACGCACACA ACAACAACGT CTCCCTGGTGT GTGCGTGTGT GTGTGTTTCTT ACCAGCGAAAA AAAACAATGG AAAACAATGG CAATTGCACA TACCAACAAAGCG 3050 ACCAGCCCAT GTAGTTTGCA GCTTAACATT GCAAAATCC TCACAAAACCG 3050 ACCAGCCCAAA AAAACAATGG AAAACAATGC TGAAAATTCA TCCAAAAACCG 3050 ACCAGCCCAAA ACCCTTCACT GGTGAATACA ACCTTTGACT GGAAAAACC TCCAAAAACCG 3050 AAAAAAATTCC GGTGAATACA ACCTTTCACT GGAAAAACC TCCAAAAACCG 3050 TTATATCCAACA ATTAGACCCAG TTCACAAGAAC ACCTTTCACT GGAAAAACC TCCAAAAACC 3350 TAATTTGTCAC ATTATATAC AAAACCTCA TAAAACCAAACC						
CEGITETICIE TITICAGIET GAAATATCAG TICECCAGCA ACATICACCT 2250 ATTATTEGCC TAGACCAAGG CECTATGTAC AGCCICCTGA AATGATTEGC 2300 CCTATGCGG TEAGGCTTA AACGTTAATT CGATGTTTIC TATTAGTAGA 2350 AATTAATTTT TEICATAGCC TCGTTCATT AGTGTGGAAA TCCTGATAAA 2400 GGICTTTCCT GCTCATAAAA AATGAGGATG GCATCTATG AAGGAAACAT 2450 TGATTCIGGA GCATTTTTT TTTTCCTCTC GIGTTCTTCA GCTTTTCCCC 2500 ATTACACACA GAGTATAATA GCCCCCAAAG TCCTGGTGC TCACTTTTCT TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTGGTCC TCACTTTTCT CCACACTAACC TCCTGTGCG TGCACTAAC TTATTGGCC TAGACCATAGG ACTTGGAGAA ATGCTTTAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 2750 ACTTGGAGAA ATGCTTTAG GCTTAGTTAG GACCAGAAACT TATTTTTTAT 27750 TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACCAAACT TATTTTTTAT 27750 TTCCACACCC ACACGTATC CCCACCCCA GCACTTCAG GTCACGTTCT 2850 AACCACCACAC ACACGTATC CCCACCCCA GCACTTCAG GTCACGTCCT 2850 CCCACCTGCTC AGGACGCAGA GGATGACGGA GCACTCACG GTCACGTCCT 2850 CCCACCTGCTC AGGACGCAGA GGATGACGGA GCACTCACG GTCACGTCCT 2850 CCCACCTGCTC AGGACGCAGA GGATGACGGA GCACTCACG GTCACGTCCT 2850 CCCACCTGCTC AGGACGCAGA GGATGACGGA GCATTCACG GTCACGTCCT 2850 CCCACACTGCTC AGGACGCAGA GGATGACGGA GCATTCACG GTCACGTCCT 2850 CCGCAAACGCA ACACACATCC CTCCTGGTGT GTCGGTGTGT GTCGTTCCT 2950 CGGAAACGCAA AAACAATCG AAACAATCC TCCCTGGTGT GTCGGTGTGT GTCGTTCCT 2950 CGGAAACGCAA AAACAATCG AAACAATCC TCCACGTGTG GTCACGTTCAC TCACCTATCAT 3000 AAAACATTCC GCCATTCACTT AAACACCC TCACAAATCC TCACCTATCAT 3100 CCTTAAACTCC GCCTTTACCTT AAAACCC GCAATTGTCC TCCACTAGCAC ACCTTTCTAC 3150 CCTTAAAACCC AAAACACC CCTTTCCTT TCCTTCACCT TCCTTAGTAC 3350 CCTTAAAACCC AAAACACC TTCACACTAC ACCTTTCTAC 3350 CCCCCAAAACG AATGCACCAG TTCACACGAC TCCAAAAACC 33550 CCCCAAAAAC AACCTTCACC TTACCTTT ACTTCATCAT TTTTTATATCG 3400 CCCTAAAATG CTTAAAACCC CCCTTTACCT ACCTTTATCA TTTTTATATCG 3400 CCTTACACACA ATCCCCCAC TCCAAAAACC TCCACACC TCCAAAAACC 33550 CCCTAAAATG CTTAAAACCC CCCTTTACCT ACCTTTCAT TTTTTATATCG 3400 CCCTAAAATCC TAAATCACGA TACTCCCATA CACGTACCT TCCAAAAACC 33550 CCCTAAAATCC TTTTCCTACCT AAAATCACGA TACTCCCATA ATCCTTTTAT TTTTTATATCG 3400 CCCTAAAAACCA TTTTTCCACC GCCCTTTACC ACCTTTTATA ATCCTTTTAT TTTTTATATCG						
ATTATIGGC TAGACCAAGG CGCTATGTAC AGCCICCTGA AATGATTGGG 2300 CCTATGCGGG TGAGTGCTTA AACGTTAATT CGATGTTTIC TATTAGTAGA 2350 AATTAATTTT TGTGATAGCG TCGTTGCATT AGTGTGGAAA TGCTGATAAA 2400 GGICTITCCT GCTCATAAAA AATGAGGATG GCATCTCATG AAGGAAACAT 2450 TGATTCTGGA GGATTTTTT TTTTCCTCTC GTGTTCTTCA GCTTTTGCC 2500 ATGACTTCTT TCTCCGCCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2550 ATGACTTCTT TCTCCGCCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2550 TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGCG TCGAGTAACC TTATTGGCCA TAGACCATAG 2650 AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 27700 CACAGTAACA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 27700 TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACGAAACT TATTTTTTAT 2750 AACCACCACA CACGTAAGAA GGGGAACCAG CAACTCAACG TCACGATCCT 2850 ACCACCTGCTC AGGAGGAGAA GGGGAACCAG CAACTCAACG TCACGATCCT 2850 CGCAACAGCAA AGACAACGT CTCCTGGTGT GTGCGTGTGT GTGTTCTCT GCGCAAAGCGA AGACAACGT CTCCTGGTGT GTGCGTGTGT GTGTTCTCT GCGCAAAGCGA AGACAACGT CTCCTGGTGT GTGCGTGTGT GTGTTCTCT GCGCAAAAGCAA AAACAATCG AAACAATCC TGCAAGTGCA TAGCAGCAAAG ACCACGAAAG AAAACAATCG AAACAATCC TGCAAATTCC TGCAAAAACG 3000 ACCACGGAAAG AAAACAATCG AAACAATCC TGCAAATTCC TGCAAAAACG 3000 ACCACGGAAAG AAAACAATCG AAACAATCC TGCAAATTCC TGCAAAAACG 3000 ACCACGGAAAG AAACAATCG GCTTAGCTTA GCCAAATCCC TCACTATGAT 3100 AAAAGTTCTC GACTTTACA ATGAGACAAT GCAAGATCC TCACTATGAT 3100 AAAAGTTCTC GACTTTACA ATGAGACAAT GCAAGAACG AAGTTGTTCT GCAAAAACC TTCACTATCAT GTATGTTCTC GCAAAAACC TCACTATGAT 3220 GGTTTGTGTC TTCCTATCAT GTATGTTCCT GCAAAAACC GCATTGTTCT 3250 GCTGAAAATG CTTAAAACC AAAAGCCTTT ACTGTAAGGT ACCTTTATAT TTTTATATTGG 3400 TACGTTTGAT AAAACCTCA TAAAACCCA TACTGCAACA TCCACTATAT TTTTTATATTGG 3400 TACGTTTGAT AAAACCTCA TAAATCACA TACTTCATAT TTTTTATATTGG 3400 TACGTTTGAT AAAACCTCCA TAAATCACA TACTTCATAT TTTTTATATTGG 3400 TACGTTTGAT AAAACCTCCA TAAATCACGA TACTTCCATA CACGTACAAA CC TTCAAAAAACC CCCCTTTTACG ACGTGCATTAA ATCCTTTTAT 3500 TAACTTTCAT AAAACCTCCA TAAATCACGA TACTTCCATA TTTTTATATCG 3400 TACGTTTAGAAAA CCAAAAAACG ATAATCACGA TACTTCCATA TTTTTTTATATCG 3600 TAATTTTCCACGCA TAAAA	CCGITCICIC	TTTTCAGIGT	GAAATATGAG	TIGGOGAGGA	AGATOGACCT	
CCTATGOGG TGAGTGCTTA AACGTTAATT CGATGTTTC TATTAGTAGA 2350 AATTAATTTT TGTGATAGG TCGTTGCATT AGTGTGGAAA TGCTGATAAA 2400 GGTCTTTCCT GCTCATAAAA AATCAGGATG GCATCTCATG AAGGAACAT 2450 TGATTCTGGA GGATTTTTT TTTTCCTCTC GTGTTCTCA GCTTTTGCCC 2500 ATGACTTCTT TCTCCGGCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2550 TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTGG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGG TGGAGTAACC TTATTGGGCA TAGAGCATAG 2650 AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT ACCTATGTAT 2770 TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACGAAACT TATTTTTTAT 2750 ACACGCACC ACCGTAATC CCCAGCCCGA GCAGTTCAGT GATGAAGTGC 2800 ACCACCACAC ACCTGAAGAA GCGAACACAG GCAGTTCAGT GATGAAGTGC 2800 GCGCAAAGGGA AGAAGAACAT CTCTCGTGTGT GTGCGTGTGT GTGCGTGTGT GTGCGTGTAG 2900 GCGCAAAGGGA AGAAGAACAT CTCTCGTGTGT GTGCGTGTGT GTGCGTGTAC 2950 GGGAAAGGGA AGAAGAACAT CTCTCGTGTGT GTGCGTGTGT GTGCGTGTACAC TAGGAGCAAC 2900 ACCACCACAC ACCTGAAGA GCGAACACAG CAATCTCACA TAGGAGCAAC 2900 GGGAAAGGGA AGAACAACGT CTCCTGGTGT GTGCGTGTGT GTGTGTTCGT 2950 GGGGAAAGGGA AGAACAACGT CTCCTGGTGT GTGCGTGTGT GTGTGTTCGT 2950 ACCACCACAAA AAACAATGG AAACAATGC CTCTGGTGT GTGCGTGTGT GTGGTGTCGT GTGGTTTCGT 2950 ACCACCAAAA AAACAATGG AAACAATGC TGAAAATTCAC TGCAAAAACC 3050 ACGAGCACAAA AAACAATGG AAACAATGC TGAAAATCC TCACTATGAT 3100 AAAAGTTCTC GACTTTATGA ATGAGAGAAT GCGAAATCC TCACTATGAT 3100 AAAAGTTCTC GACTTTATGA ATGAGAGAAT GCGAAATCC TCACTATGAT 3220 GGTTTGTGTC TTCCTATCAT GTAGTTTCCT GGAAAACCC TCACTATGAT 3250 GGTTTGTGTC TTCCTATCAT GTATGTTCCT GGAAAACCC TCACTATGTAC 3300 GGTTTGTGTC TTCCTATCAT GTATGTTCCT GGAAAACCC TCCACAAAACC 3350 GGGTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGT ACCTTAGTAC 3300 TAACCTTTGAT AAAACCTCA TAAATCACA TCCTTAGTAC TTTTTTATTTTT 3250 GCGTAAAAATG CTTTAAAACTC AAAAGGCTTT ACTGTAAATT TTTTTATTTTT 33500 GGTTTAGAAAA GCAAAAACCG TTCAGAGAGA GAGCCATTTA AAACCTCTAA AAACCTTTATTT TCATGAGGAG CAGGACCAC TCCAAAAACC 3350 TTAATTCTCC GACATAAAAT GTTTTCAGA GGGTGATTAA ATCCTTTTTAT 3500 GGTTTAGAAAA GCAAAAACCG ATTATACCTG AGATTAACGT GACATGGAAA 3550 GTTTAGAAAAACC GAAAAAACGG ATTATACCTT AAAGTTCCTA GAAATCCTAC GACATGCAAA 3650 TTAATTCTCC GACATTAAAAT GTTT						
AATTAATTIT TGTGATAGOG TOGTTGCATT AGTGTGGAAA TGCTGATAAA 2400 GGTCITTCCT GCTCATAAAA AATGAGGATG GCATCICATG AAGGAAACAT 2450 TGATTCTGGA GGATTITTT TITTCCTCTC GTGTTCTCA GCTTTTGCCC 2500 ATGACTTCTT TCTCCGGCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2550 TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACC TCCTGTGCG TGGAGTAACC TTATTGCGCA TAGAGCATAG 2650 AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 2700 TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACCAAAACT TATTTTTTAT 2750 TTCCACACCC ACACGTATC CCCACCCCGA GCAGTTCAGT GATGAAGTGC 2800 AACCACCACAC ACCTGAACAA GGGGAACCAG CAACTCAACG TCACGATCCT 2850 CACACTGCTC ACCAGGCACA GGATGAGGGA GCATTCTCCAG GTCAAGGTGA 2900 CGGAAAGGGA ACAACAACGT CTCCTGGTGT GTGCGTGTGT GTGTGTTCGT 2950 GTGTGTGTGT GCACGTGTG GTGGTTTAGC CATTGTCACA TAGCAGCAAG 3000 ACGAGCAAAG AAACAATCG AAACAATCC TGCAGGTGA CATTGTCACA TAGCAGCAAG 3000 ACGAGCAAAG AAACAATCG AAACAATCC TGAAAATCCC TGAAAATCAC TGCAAAACCC 3050 ACGAGCCTACT GTAGTTTCCA CCTTAGCTTA GGCAAAATCCC TGCAAAATCCC TGCAAAATCCC TGCAAAATCCC TGCAAAATCCC TGCAAAATCCC TGCACGTGTG TCCTATGAT 3100 AAAACATTCTC GACTTTAACA ATGAGAGAAT GGACGTCCCA GCATTGTGTG 3150 TTATCCAACA ACCCTTCACT GTAAATCA ACCTTTGTAC TGTGTTCTTAA 3200 CGTTAAAATC CTTAAAACTC AAAAGCCTTT ACTGTAAGGT ACCTTTAGTAC 3300 TGACCCAACA ATGACCCAG TTCCAACAGC CACGACCACC TCCAAAAACC 3350 TGACCCAACA ATGACCCAG TTCCAACGAG CACGACCACC TCCAAAAACC 3350 TGACCCAACA ATGACCCAG TTCAACACC CACGTTCCTT TCATTCATAC TTTTTTATTTCCAC CCGTTTACTT TCATTCATT TTTTTTATTTCC 3400 TTACGTTTGAT AAAACCTCCA TAAAACCC AAAACCCAG TTCCAAAAACC 3350 TGACCCAACAA ATGACCCAG TAAAACCC AAAACCCAG TTCCAAAAACC 3350 TGACCTTTCAT AAAACCTCCA TAAAACCC AAAACCCAG TACCTACAAA ACCTTTTTTT TCATTCATCT TCATTCAT						
GETCHTICT GCTCATAAAA AATGAGGATG GCATCTCATG AAGGAAACAT TGATTCTGGA GGATTTTTT TTTTCCTCTC GTGTTCTTCA GCTTTTGCCC 2500 ATGACTTCTT TCTCCGGCTT TGTTTGTTAA TGACAGATTG TACACATGTA ATGACTACACA GAGTATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGGG TGGAGTAACC TTATTGGCCA TAGAGCATAG AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT ACCTATGTAT TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACCAAACT TATTTTTTAT TTTCCACACCC ACACGTATTC CCCAGCCCGA GCAGTTCAGT GATGAAGGTGG AACCAGCAAC ACCTGAAGAA GGGGAAACCAG CAACTCAACG TCAGCATCCT AGCACGCACA ACCTGAAGAA GGGGAAACCAG CAACTCAACG TCAGCATCCT AGCACGCACAC ACCTGAAGAA GGGGAAACCAG CAACTCAACG TCAGCATCCT AGCAACGCA ACACAGAACA GGGGAAACCAG CAACTCAACG TCAGCATCCT AGCAACGCA AGAACAACCT CTCCTGGTGT GTGCGTGTGT GTGTGTTCGT GGGAAAGCGA AGAACAACCT CTCCTGGTGT GTGCGTGTGT GTGTGTTCGT ACGACCAAAG AAACAATCG TCACTGGTGT GTGCGTGTGT GTGTGTTCGT AAAACAATCC TCACTGTTAG CATTGTCACA TAGCACCAAG ACGACCAAAA AAACAATCG AAACAATCC TCAAAATCCC TCACAGACCAAG AAACAATCC TCACAGTTAA ACCCTTACCTTA GCCAAATCCC TCACATACGT TTAATCCAACA ACCCTTCACT GGTGAATACA ACATTTGTAC TGTGTTCTAA AAAAGTTCTC GACTTTATCA ATGAGACCAAT GCACGACACC TCCAAAAACC GGTTTGGTC TTCCTATCAT GTATGTTCCT GCAAAACAACG AAGTGATTTT 3250 CGTGAAAATG CTTAAAACTC AAAAGCCTTT ACTGTAACGT ACCTTAGTAC 3300 TGACCCAACA ATGACCCAG TTCAGAGGAG CAGGACCAC TCCAAAAACC GGTTTAGTAT AAAACCCAG TTCAGAGGAG CAGGACCAC TCCAAAAACC GAGTCCCTCA ATGTTGCCC CCGTTTCCTT TCATTCATAT TTTTTATATGG 3400 TACGTTTGAT AAAACCCAG TCAGAGCAC TCCAAAAACC 3350 TACGTTTGAT AAAACCTCCA TAAATCACGA TACTCCCATA CACGTACCTG 3450 TACGTTTGAT AAAACCTCCA TAAATCACGA TACTCCCATA CACGTACCTG 3450 TTACTTTCTAC GCCCTTTTAGC ACGTTTACAT TTTTTTATATTCG 3400 TTACTTTCAC GCCTTTTAGC ACGTTTACAT TTTTTTTATATTCG 3400 TTACTTTCAC GCCCTTTTAGC ACGTTTACAT TTTTTTTTTT						
TEATTCTEGA GEATTTTTT TTTTCCTCTC GIGITCTTCA GCTTTTGCCC 2500 ATGACTICTT TCTCCGCTT TGTTTGTTAA TGACAGATTG TACACATGTA 2550 TTCCAACACA GAGTATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGCG TGGAGTAACC TTATTGGCAATAG 2650 AGTTGGAGAA ATGTCTTTAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 2700 TCTGTGTATA TATGTAAAAT TTTTGTATCAA TAACGAAACT TATTTTTTAT 2750 TTGCACACCC ACACGTATTC CCCAGCCCGA GCAGTTCAGT GATGAAGTGG 2800 AACCAGCAAC ACCTGAAGAA GCGGAACCAG CCAGTTCAGT GATGAAGTGG 2800 AACCAGCACC ACACGTATTC CCCAGCCCGA GCAGTTCAGT GATGAAGTGG 2800 GCGAAAGCGA AGAACAACT CTGCTGGTGT GTGCGTGTT GTGAGTTCGT 2950 GCGCAAAGCGA AGAACAACCT CTGCTGGTGT GTGCGTGTGT GTGGTGTTCGT 2950 GGGAAAGCGA AGAACAACCT CTGCTGGTGT GTGCGTGTGT GTGGTGTTCGT 2950 GGGAAAGCGA AGAACAACCT CTGCTGGTGT GTGCGTGTGT GTGGTGTTCGT 2950 ACGAGCCAAAG AAAACAATCG AAACAATCC TGCAAAATCCA TAGCAACGAAG 3000 ACGAGCCAAAG AAAACAATCG AAACAATCC TGCAAATCCA TGCGAAAACCG 3050 ACGAGCCAAAG AAAACAATCG GCTTAGCTTA GCCAAATCCA TGCGAAAACCG 3050 AAAAAGTTCTC GACTTTATGA ATGAGACAAT GCAGGTGCCA GGATTGTGT 3150 TTATCCAAGA ACCCTTGACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200 GGTTTGTGTC TTCCTATCAT GTAGTTTCCT GGAAAAACC ACATTTGTAC TGTGTTCTAA 3250 GCGTGAAAATG CTTAAAACTC AAAAGCCTTT ACTGTAAGGT ACGTTAGTAC 3350 TGACCCAACA ATGGACCCAG TTCAGAGGAG CAGGACCAGC TCCAAAAACC 3350 TGACCCAACA ATGGTCCCAG TTCAGAGGAG CAGGACCAGC TCCAAAAACC 3350 TGACCCAACA ATGGTCCAG TAAATCCA TACTGCATAA CAGGTAGCTG 3450 TAACGTTTGAT AAAACCTCGA TAAATCACGA TACTGCCATA CAGGTAGCTG 3450 TAACGTTTGAT AAAACCTCGA TAAATCACGA TACTGCCATA CAGGTAGCTG 3450 TGACTTAGAAAA GCAAAAACG TAAATCACG AGCTTTAAA ATCCTTTTAT 3500 GGTTTAGAAAA GCAAAAACG AATTACCTG AGATTAACGT GAGATCCAAA 3550 TTAATTTCTCAC GCCCTTTTAGG AGGTTTAAAACCT GAGATCCAAA 3550 TTAATTTCTCAC GCCCTTTTAGG AGGTTTAAAACCT GAGATCCAAA 3550 TTAATTTCTCAC GCCCTTTTTAGG AGGTTTAA ATCCTTTTAT 3500 GGTTAGAAAAA GCAAAAAACG AATTATCCTG AGATTAACGT GAGATCCAAA 3550 TTAATTTTCCAC GCCCTTTTTAAAACTT GAGATTAAACT GAGATCAAAA 3550 TTAATTTCTCAC GACATAAAAC GTTTTTGAAAAC GAATCAACA AATCCTTTTAT 3500	CCICITICCT	GCTCATAAAA	AATGAGGATG	GCATCTCATG	AAGGAAACAT	
ATGACITCIT TCICCGCCIT TGITTGITTAA TGACAGATTG TACACATGIA 2550 TTCCAACACA GAGIATAATA GCCCCCAAAG TCCTCGTGCG TCACTTTTCT 2600 CACAGTAACC TCCCTGTGCG TGGAGTAACC TTATTGGCCA TAGAGCATAG 2650 AGTTGGAGAA ATGICTITAG GCTTAGTTAG GACCAGAAAT AGCTATGTAT 2700 TCTGTGTATA TATGTAAAAT TTTGTATCAA TAACCAAACT TATTTTTTAT 2750 TTCCACACCC ACACGIATTC CCCAGCCCGA GCAGTTCAGT GATGAAGTGG 2800 AACCAGCAAC ACCTGAAGAA GGGGAACCAG CAACTCAACG TCAACGTCCT 2850 GCACTGCTC AGGAGGCAGA GGATGAGGGA GCATTGCCAG GTCAAGGTGA 2900 GGGAAAGGCA ACAAGAACGT CTGCTGGTGT GTGGTGTTGGT 2950 GGGGAAAGGCA ACAAGAACGT CTGCTGGTGT GTGGTGTTCGT 2950 GGGAAAGGCA ACAAGAACGT CTGCTGGTGT GTGGTGTTCGT 2950 GGGAAAGCAA AAAACAATGG AAAGAATGC TGAAAATTGAC TGGAAAAAGCG 3050 ACGAGCAAAG AAAACAATGG ACAAGAATGC TGAAAATTGAC TGGAAAAAGCG 3050 ACGAGCAAAG AAAACAATGG ACAAGAATGC TGAAAATTGAC TGGAAAAAGCG 3050 ACGAGCAAAA ACCCTTGACT GGTGAATACA ACATTTGTAC TGGTTCTAA AAAAGTTCTC GACTTTATGA ATGAGAGAAT GGAGGTGCCA GCATTGTGTAA ACCTTGACT GTAGTTTCCT GGTGAATACA ACATTTGTAC TGGTTCTAAA ACCTTGACT TTCCTATCAT GTAGTTTCCT GGAAACAAGG AAGTGATTTT 3250 GGTGAAAATG CTTAAAACTC AAAAGCCTTT ACTGTAAGGT ACCTTAGTAC 3350 GGAGTOCCTGA ATGTGCCC CCGTTTCCTT TGATTGATAT TTTTTATTTGG 3400 TACGTTTGAT AAAACCTCA TAAAACCC AAAAGCCATA CAGGTACCTG TCCAAAAACC 3350 TAACTTTGAT AAAACCTCA TAAAACCC AAAACCAATACA ACGTTAGTAT TTTTTATTTGG 3400 TACGTTTGAT AAAACCTCAAAAACCC AAAATCACA ACGTTAGTAT TTTTTATTTGG 3400 TACGTTTGAT AAAACCTCAAAAACCG ATTAAACACA ACGTTAGTATA ACCCTTTTATT TTTTTATTTTTTATTTTTTTTTT	TGATTCTGGA	GGATTTTTTT	TITTCCTCTC	GIGITCITCA	GCTTTTGCCC	
CACAGTAACC TCCCTGTGGG TGCAGTAACC TTATTGGCA AGTTGCAGAA ATGICTTTAG GCTTAGTTAG GACCACAAAT AGCTATGTAT 2700 TCTGTGTATA TATGTAAAAT TTTTGTATCAA TAACGAAACT TATTTTTTAT 2750 TTGCACACCC ACACGTATTC CCCACCCCGA GCAGTTCAGT GATGAAGTGC ACACGTATTC CCCACCCCGA GCAGTTCAGT GATGAAGTGC ACACGTATTC CCCACCCCGA GCAGTTCAGT GATGAAGTGC ACACGTACT AGCAGCAGCA ACACGTACT AGCAGCAGCA GCAGTGCGC ACACGTGCT ACGAGCAGCA GCAGCAGCA GCAGTGCGC GCAGTGCGC GCAGTGCGC ACACGTGGCI GTGTGTTAGC GCAGTGGTG GTGTGTTCGT GTGCGTGTGT GTGCGTGTGT GTGCGTGTGCT GTGCGTGTGCT GTGCGTGTCT ACCACACAAA ACACTTCACC ACACGTACC ACACGCAC ACACGTACC ACACGC ACACGCAC ACACCAC ACACCAC ACACCAC ACACCAC ACACCAC						
AGTIGGAGAA ATGICITIAG GCITAGITAG GACCAGAAAT AGCIATGIAT 2750 TCIGIGTATA TATGIAAAAT TITGIATCAA TAACCAAACT TATTITITAT 2750 TTIGCACACCC ACACGIATIC COCAGCOCGA GCAGTICAGT GATGAAGIGG 2800 AACCAGCAAC ACCTGAAGAA GCOGAACCAG CAACTCAACG TCAGGATCCT 2850 GCAGCIGCIC AGGAGGAGA GGATGAGGGA GCATCICAG GICAACGIGA 2900 GGGAAAGGGA AGAAGAACGT CIGCIGGIGI GIGCGIGIGI GIGIGITCGT 2950 GTGIGTGIGT GCACGIGIGI GIGIGITAGG CATTGICACA TAGGAGGAAG 3000 AGGAGGAAAG AAACAATGG AAAGAATGCC TGAAAATTGAC TGGAAAAGCG 3050 AGGAGGAAAG AAACAATGG AAAGAATGCC TGAAAATCCC TCACTATGAT 3100 AAAAGTICIC GACTITATGA ATGAGAGAAT GCAAGATCCC TCACTATGAT 3150 TTATCCAAGA ACCCTTGACT GGIGAATACA ACATTTGIAC TGIGTICIAA 3200 GGTTGGIGC TTCCTATCAT GGIGAATACA ACATTTGIAC TGIGTICIAA 3200 GGTTGGIC TTCCTATCAT GTAGATCCT ACTGIAAGGT AGCTTAGTAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGCACCAC TCCAAAAACC 3350 GAGTCCCTGA ATGTCCCCC COGITTCCTT TCATTGATAT TTTTATATGG 3400 TACGITTGAT AAAACCIGGA TAAATCAGCA TACTGCATTA CAGGIACCTG 3450 GGTTAGGAT TTTTCTCACC GCCCTTTAGC ACGTCATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 GGTTAGAAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 GGTTAGAAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 GGTTAGAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 GGTTAGAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3600 GTCATGGATT ATTCCACGGA TCCACTGTTA AAAGTTCCTA GAATCCTACTA GAGTTCGACT 3600						2600
TCIGIGIATA TAIGIAAAAT TITIGIATCAA TAACCAAACT TAITITITAT 2750 TTGCACACCC ACACGIATIC COCAGCOCGA GCAGTICAGT GATGAAGIGG 2800 AACCAGCAAC ACCIGAAGAA GGGGAACCAG CAACTCAACG TCAGGATCCT 2850 GCAGCTGCTC AGGAGGAGA GGATGAGGGA GCATCIGCAG GICAAGGIGA 2900 GGGAAAGGCA ACAAGAACGT CIGCIGGIGT GIGCGIGIGT GIGIGITOGT 2950 GTGIGIGIGT GCACGIGIGT GIGIGITAGG CATTGICACA TAGGAGGAAG 3000 ACGAGGAAAG AAAACAATGG AAAGAATGCC TGAAAATTCAC TGGAAAAGCG 3050 ACGAGGAAAG AAAACAATGG AAAGAATGCC TGAAAATTCAC TGCAAAAACGG 3050 ACGAGGCTAT GTAGTTTGCA GCTTAACTTA GCACAAATCCC TCACTATGAT 3100 AAAAGTICIC GACTTTATGA ATGAGAGAAT GCAGGIGCCA GCATTGIGIG 3150 TTATCCAAGA ACCCTTGACT GGIGAATACA ACATTTGTAC TGIGTTCTAA 3200 GGTTTGIGIC TTCCTATCAT GTATGTICCT GGAAACAACG AAGTGATTTT 3250 GCTGAAAATG CTTAAAACTC AAAAGCCTT ACTGTAAGGT ACCTTAGTAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGCACACC TCCAAAAACC 3350 TGACCCAAGA ATGAGCCCA TAAATCAGCA TACTGCAATA CAGGIAGCTG 3450 GAGTCGTGAAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA GCTTTAGTACT TTTTTCTCACC GCCCTTTAGC ACGTCATTAA ATCCTTTTAT 3500 GGTTTAGTGAT TTTTTCTCACC GCCCTTTAGC ACGTCATTAA ATCCTTTTAT 3500 GGTTTAGAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTCAAAG GAACCATTTA TGIAACCGAG 3600 GTCATGGATT ATTCCACGGA TCCCACTTA CAGGTACGAAA 3550 GGTTAGAAAAA CCAAAAACCG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 GGTTAGAAAAACCT GAGATAAAAT GTTTTCAAAAG GAACCATTTA TGIAACCGAG 3600 GTCATGGAT ATTCCACGGA TCCACCTGTTA AAAGTTCCTA GAATCCTACTAACT GAGATCGACT 3650						2650
TIGCACACC ACACGIATIC COCAGCOCGA GCACITCAGI GATGAAGIGG 2800 AACCAGCAAC ACCIGAAGAA GGGGAACCAG CAACTCAACG TCAGGATCCT 2850 GCAGCIGCIC AGGAGGGAG GGATGAGGGA GCATCIGCAG GICAAGGIGA 2900 GGGAAAGGGA AGAACAACGI CIGCIGGIGI GIGCGIGIGI GIGIGIITOGI 2950 GIGIGIGIGIGI GCACGIGIGI GIGIGIITAGG CATTGICACA TAGGAGGAAG 3000 AGGAGGAAAG AAAACAATGG AAAGAATGCC TGAAAATGAC TCGCAAAAGCG 3050 AGGAGGCTAT GTAGITIGCA GCITACCITIA GGCAAAATCCC TCACIATGAT 3100 AAAAGITCIC GACTITATGA ATGAGAGAAT GGAGGIGCCA GGATTGIGIG 3150 TIATCCAAGA ACCCTIGACT GGIGAATACA ACATTIGIAC TGIGITCITAA 3200 GGTTTGIGIC TTCCTATCAT GTATGITCCT GGAAACAACG AAGTGATTIT 3250 GCTGAAAATG CTTAAAACIC AAAAGGCTIT ACTGIAAGGT AGCTTAGIAC 3300 TGACCCAAGA ATAGACCAG TTCAGAGGAG CAGGACCACC TCCAAAAACC 3350 GAGTCGCTGA ATGITGGCCC CCGITTCCTT TGATTGATAT TITTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGIAGCTG 3450 GGTTAGGATA AAAACCTGGA TAAATGAGGA AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAAACG AATTATCCTG GAGATAAAAC GAGATGGAAA 3550 GGTTAGAAAAACG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAAACG AATTATCCTG AGATTAACGT GAGATGGAAA 3600 GGTTAGGAATAAAAT GTTTTGAAAG GAACCATTTA TGIAACGGAG 3600 GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACTT ATTCCAGGGATTAAACTTCACTT GAACGTACTAC GAATCTGACTT ATTCCACTA GAATCTGACTT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT						2700
AACCAGCAAC ACCTGAAGAA GGGGAACCAG CAACTCAACG TCAGGATCCT 2850 GCAGCTGCTC AGGAGGAGA GGATGAGGGA GCATCTGCAG GTCAAGGTGA 2900 GGGAAAGGGA AGAAGAACGT CTGCTGGTGT GTGCGTGTT GTGTGTTCGT 2950 GTGTGTGTGT GCACGTGTGT GTGTGTTAGG CATTGTCACA TAGGAGGAAG 3000 AGGAGGAAAG AAAACAATGG AAAGAATGCC TGAAATTCAC TGGAAAAGCC 3050 AGGAGGCTAT GTAGTTTGCA GCTTAGCTTA GGCAAATCCC TCACTATGAT 3100 AAAAGTTCTC GACTTTATGA ATGAGAGAAT GGAGGTGCAA GGATTGTGTG TTATCCAAGA ACCCTTGACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200 GGTTTGTGTC TTCCTATCAT GTATGTTGCT GGAAAGAAGG AAGTGATTTT 3250 GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT AGCTTAGTAC 3300 TGACCCAAGA ATGAGACCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTCGCTGA ATGTTGCCC CCGTTTCCTT TGATTGATAT TTTTATATGG 3400 TACGTTTGAT AAAACCTGGA TAAATCAGGA TACTGCCATA CAGGTACCTG 3450 GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAAA GCAAAAACGG AATTATCCTG AGATTAACT GAGATGCAAA 3550 GGTTAGAAAAA GCAAAAACGG AATTATCCTG AGATTAACT GAGATGCAAA 3550 GGTTAGAAAAA GCAAAAACGG AATTATCCTG AGATTAACT GAGATCCAAAA 3550 GGTTAGCAATA ATCCTGCAATAAAT GTTTTCAAAG GAACCATTTA TGTAACCGAG 3600 GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCTA TGTAACCGAG GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCTA TGTAACCGAG GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT AAATTTCTCC GAGATAAAAT GTTTTCAAAAG GAACCATTTA TGTAACCGAG GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT GACATCGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT GCACATGGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT GACATCGATT ATTCCAGGGA TCCACTGTTA AAAGGTTCCTA GAATCTGACT						2750
GCAGCTGCTC AGGAGGCAGA GGATGAGGGA GCATCTGCAG GTCAAGGTGA 2900 GGGAAAGGGA AGAAGAACGT CTGCTGGTGT GTGCGTGTGT GTGTGTTCGT 2950 GTGTGTGTGT GCACGTGTGT GTGTGTTAGG CATTGTCACA TAGCAGGAAG 3000 AGGAGGAAAG AAAACAATGG AAAGAATGCC TGAAATTGAC TGGAAAAGCG 3050 AGGAGCCTAT GTAGTTTGCA GCTTACCTTA GGCAAATCCC TCACTATCAT 3100 AAAAGTTCTC GACTTTATGA ATGAGAGAAT GGAGGTGCCA GCATTGTGTA 3200 GTTATCCAAGA ACCCTTCACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200 GGTTTGTGTC TTCCTATCAT GTATGTTGCT GGAAACAAGG AAGTGATTTT 3250 GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT AGCTTAGTAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTCGCTGA ATGTTGGCCC CCGTTTCCTT TGATTGATAT TTTTTATATGG 3400 TACGTTTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG ACGTCATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCCAAAAACCG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAA GCCAAAAACCG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAA GCCAAAAACCG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAAA GCCAAAAACCG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGAAAAA GCCAAAAACCG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 GGTTAGGATT ATTCCAGGGA TGCACTGTTA AAAGGTTCCTA GAATCTGACT						2800
GGGAAAGGGA AGAAGAACST CTGCTGSTGT GTGCGTGTGT GTGTGTTCGT 2950 GTGTGTGTGT GCACGTGTGT GTGTGTTAGG CATTGTCACA TAGGAGGAAG 3000 AGGAGGAAAG AAAACAATGG AAAGAATGCC TGAAAATGAC TGGAAAAGCG 3050 AGGAGCCTAT GTAGTTTGCA GCTTAGCTTA GGCAAAATCCC TCACTATGAT 3100 AAAAGTTCTC GACTTTATGA ATGAGAGAAT GGAGGTGCCA GGATTGTGTG 3150 TTATCCAAGA ACCCTTGACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200 GGTTTGTGTC TTCCTATCAT GTATGTTGCT GGAAAGAAGG AAGTGATTTT 3250 GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT AGCTTAGTAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTCGCTGA ATGTTGCCCC CCGTTTCCTT TGATTGATAT TTTTATATGG 3400 TACGTTTGAT AAAAGCTGGA TAAAATGAGGA TACTGCCATA CAGGTACCTG 3450 GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATCGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAAGCATTTA TGTAACGAG 3600 GTCATGGATT ATTCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT						2850
GIGIGIGIGI GCACGIGIGI GIGIGITIAGG CATTGICACA TAGGAGGAAG  ACGAGGAAG AAAACAATGG AAAGAATGCC TGAAATTGAC TGGAAAAGCG  AGGAGGCTAT GTAGTTIGCA GCTTAGCTTA GGCAAATCCC TCACIATGAT  AAAAGITCIC GACTTTATGA ATGAGAGAAT GCAGGIGCCA GCATTGIGIG  TTATCCAAGA ACCCTTGACT GGIGAATACA ACATTTGIAC TGIGITCTAA  3200  GGTTIGIGIC TTCCTATCAT GTATGTTGCT GGAAAGAAGG AAGTGATTTT  3250  GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT ACCTTAGTAC  TGACCCAAGA ATGAGACCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC  3350  GAGTOGCTGA ATGTTGGCCC CCGTTTCCTT TGATTGATAT TTTTATATGG  3400  TACGTTTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG  3450  GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTCATTAA ATCCTTTTAT  3500  GGTTAGAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATCGAAA  GCATAGAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATCGAAA  3550  TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGIAACGGAG  GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGTTCCTA GAATCTGACT  GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGTTCCTA GAATCTGACT  GTCATGGATT ATTCCAGGGA TCCACTGTTA AAAGTTCCTA GAATCTGACT  GCACACAATTAACTA GAATCTGACT GAGATCTGACT  3650						2900
AGGAGGAAAG AAAACAATGG AAAGAATGCC TGAAAATTGAC TGGAAAAGCG 3050 AGGAGGCTAT GTAGITIGCA GCTIAGCTTA GGCAAATCCC TCACIATGAT 3100 AAAAGITCIC GACITTATGA ATGAGAGAAT GGAGGIGCCA GGATTGIGIG 3150 TTATCCAAGA ACCCITGACT GGIGAATACA ACATTTGIAC TGIGTICIAA 3200 GGITIGIGIC TTCCIATCAT GIATGITGCT GGAAAGAAGG AAGTGATTTT 3250 GCIGAAAATG CITAAAACIC AAAAGGCITT ACTGIAAGGT AGCTTAGIAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGICGCTGA ATGITGGCCC CCGITTCCTT TGATTGATAT TTTTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAAATGAGGA TACTGCCATA CAGGIAGCTG 3450 GTTIAGTGAT TTTTCICAGC GGCCTTTAGG AGGIGATTAA ATCCITTTAT 3500 GGTIAGAAAA GCAAAAACG AATTATCCTG AGATTAACGT GACATGAAA 3550 TAATTTCTCC GAGATAAAAT GITTTGAAAG GAAGCATTTA TGIAACCGAG 3650 GTCATGGATT ATTCCAGGGA TCCACTGITA AAAGTICCTA GAATCTGACT 3650						2950
AGGAGGCTAT GTAGTTIGCA GCTTAGCTTA GGCAAATCCC TCACTATGAT  AAAAGTTCTC GACTTTATGA ATGAGAGAAT GGAGGTGCCA GCATTGTGTG  TTATCCAAGA ACCCTTGACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200  GGTTTGTGTC TTCCTATCAT GTATGTTGCT GGAAAGAAGG AAGTGATTTT 3250  GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT AGCTTAGTAC 3300  TGACOCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350  GAGTCGCTGA ATGTTGGCCC CCGTTTCCTT TGATTGATAT TTTTTATATGG 3400  TACGTTTGAT AAAAGCTGGA TAAAATGAGGA TACTGCCATA CAGGTAGCTG 3450  GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500  GGTTAGAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATCGAAA 3550  TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGTAACGGAG 3600  GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650						3000
AAAAGITCIC GACITTATGA ATGAGAGAAT GGAGGTGCCA GGATTGIGTG TTATCCAAGA ACCCITGACT GGIGAATACA ACATTTGIAC TGIGFICITAA 3200 GGITTGIGTC TTCCIATCAT GIATGITGCT GGAAAGAAGG AAGTGATTTT 3250 GCIGAAAATG CITAAAACTC AAAAGGCITT ACTGIAAGGT AGCTTAGIAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGCAGCAGC TCCAAAAACC 3350 GAGTOGCTGA ATGITGGCCC CCGITTCCTT TGATTGATAT TITTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGIAGCTG 3450 GTTIAGIGAT TTTTCICAGC GGCCTTTAGG AGGIGATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGIAACGGAG GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650						3050
TTATCCAAGA ACCCTTGACT GGTGAATACA ACATTTGTAC TGTGTTCTAA 3200 GGTTTGTGTC TTCCTATCAT GTATGTTGCT GGAAAGAAGG AAGTGATTTT 3250 GCTGAAAATG CTTAAAACTC AAAAGGCTTT ACTGTAAGGT AGCTTAGTAC 3300 TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTCGCTGA ATGTTGGCCC CCGTTTCCTT TGATTGATAT TTTTATATGG 3400 TACGTTTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCAAAAACGG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGTAACCGAG 3600 GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650						3100
GGITIGIGIC TICCIATCAT GIATGITGCT GGAAAGAAGG AAGIGATITT 3250 GCTGAAAATG CITAAAACTC AAAAGGCTIT ACTGIAAGGT AGCTTAGIAC 3300 TGACOCAAGA ATAGACOCAG TICAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTOGCTGA ATGITGGCCC CCGITTCCTT TGATTGATAT TITTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GTTIAGIGAT TTTTCICAGC GGCCITTAGG AGGTGATTAA ATCCITTTAT 3500 GGTTAGAAAA GCAAAAACG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGTAACGGAG 3600 GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650						3150
GCTGAAAATG CITAAAACTC AAAAGGCTTT ACTGIAAGGT AGCTTAGIAC 3300 TGACOCAAGA ATAGACCCAG TTCAGAGGAG CAGGAGCAGC TCCAAAAACC 3350 GAGTOGCTGA ATGITGGCCC COGITTCCTT TGATTGATAT TITTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GTTTAGTGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCAAAAACGG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGTAACGGAG 3600 GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650						3200
TGACCCAAGA ATAGACCCAG TTCAGAGGAG CAGCAGCAGC TCCAAAAACC 3350 GAGTOGCTGA ATGITGGCCC COGITTOCTT TGATTGATAT TTTTATATGG 3400 TACGITTGAT AAAAGCTGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GITTAGTGAT TTTTCTCAGC GGCCITTAGG AGGTGATTAA ATCCITTTAT 3500 GGTTAGAAAA GCAAAAACCG AATTATCCTG AGATTAACGT GAGATGAAA 3550 TAATTTCTCC GAGATAAAAT GITTTGAAAG GAAGCATTTA TGIAACCGAG 3600 GTCATGGATT ATTCCAGGGA TGCACTGITA AAAGTTCCTA GAATCTGACT 3650						3250
GAGTOGCTGA ATGITGGCCC COGITTCCTT TGATTGATAT TITTATATGG 3400 TACGITTGAT AAAACCIGGA TAAATGAGGA TACTGCCATA CAGGTAGCTG 3450 GITTAGIGAT TTTTCTCAGC GGCCTTTAGG AGGTGATTAA ATCCTTTTAT 3500 GGTTAGAAAA GCAAAAACGG AATTATCCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTCC GAGATAAAAT GTTTTGAAAG GAACCATTTA TGTAACGGAG 3600 GTCATGGATT ATTCCAGGGA TGCACTGTTA AAAGTTCCTA GAATCTGACT 3650	GCIGAAAATG	CITAAAACIC	AAAAGGCTTT	ACTGTAAGGT	AGCITAGIAC	3300
TACGITIGAT AAAACCIGGA TAAATGAGGA TACIGCCATA CAGGIAGCIG GITIAGIGAT TITICICAGC GGCCITTAGG AGGIGATTAA ATCCITITAT 3500 GGITAGAAAA GCAAAAACGG AATTATCCIG AGATTAACGI GAGATGGAAA 3550 TAATTICICC GAGATAAAAT GITITGAAAG GAAGCATTTA TGIAACGGAG 3600 GICATGGATT ATTCCAGGGA TGCACIGITA AAAGITCCIA GAATCIGACT 3650	TGACCCAAGA	ATAGACCCAG	TTCAGAGGAG	CAGGAGCAGC	TCCAAAAACC	3350
GITTAGIGAT TITTCICAGC GGCCITTAGG AGGIGATTAA ATCCITTIAT 3500 GGTIAGAAAA GCAAAAACGG AATTATCCIG AGATTAACGT GAGATGGAAA 3550 TAATTICICC GAGATAAAAT GITTIGAAAG GAAGCATTTA TGIAACGGAG 3600 GICATGGATT ATTCCAGGGA TGCACIGITA AAAGTICCIA GAATCTGACT 3650	GAGTOGCTGA	AIGITGGCCC	CCGITICCIT	TGATTGATAT	TTTTATATGG	3400
GGITAGAAAA GCAAAAACGG AATTATOCTG AGATTAACGT GAGATOGAAA 3550 TAATTTCTOC GAGATAAAAT GITTTGAAAG GAAGCATTTA TGIAACGGAG 3600 GTCATGGATT ATTOCAGGGA TGCACTGITA AAAGTTOCTA GAATCTGACT 3650	TACGITIGAT	AAAAGCIGGA	TAAATGAGGA	TACTGOCATA	CAGGIAGCIG	3450
GGITAGAAAA GCAAAAACGG AATTATOCTG AGATTAACGT GAGATGGAAA 3550 TAATTTCTOC GAGATAAAAT GITTTGAAAG GAAGCATTTA TGIAACGGAG 3600 GTCATGGATT ATTOCAGGGA TGCACTGITA AAAGTTOCTA GAATCTGACT 3650	GITTAGTGAT	TTTTCTCAGC	GGCCTTTAGG	AGGIGATTAA	ATCCTTTTAT	3500
TAATTICIOC GAGATAAAAT GITTIGAAAG GAAGCATITA TGIAACOGAG 3600 GICATGGATT ATTOCAGGGA TGCACIGITA AAAGTICCIA GAATCIGACT 3650	GGTTAGAAAA	GCAAAAACGG	AATTATOCTG	AGATTAACGT	GAGATGGAAA	
GICATGGATT ATTOCAGGA TGCACTGITA AAAGITOCTA GAATCTGACT 3650	TAATTICICC	GAGATAAAAT	GTTTTGAAAG	GAAGCATTTA	TGIAACGGAG	3600
CACAACAAMC COCAMMAAMM COMOMOGOGO CA COCAMMAAMA	GICAIGGAIT	ATTOCAGGGA	TGCACIGITA	AAAGITIOCIA	GAATCIGACT	
	GACAACAATG	CCCATTAATT	CCICICOCC	CACTOCCTTA	TICICAGIGC	3700

			·		
10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GGGGACAGTA	TATTTTCIGI	GATTCACAAA	CAATGITATA	TTTGGTGCTT	3750
TGTTCTTCAC	GGGGTTCATT	TATGGAATAT	TACCTITAGG	ACCTTOGGAC	3800
CTAAATATAA	CITIATTIGA	ACAAAGIGAA	GITICICITT	ACCCCGATAG	3850
GTAATGGGTG	TOGIGACIGT	AAGATTICCA	TAGICCTCAA	ATCCATCCAG	3900
CTAATCAATC	CITCAGAAAC	TGACATTGTA	ATTGEAACTG	AAATOCTACC	<b>395</b> 0
CACGIGGIAG	ACTICAGATT	TCTCACGTGA	CCCACACTCC	TGITGGIACT	4000
CTAAGGCTGA	ATATAAGCAT	TATACATGIC	CIGICGITIA	TCCTTAGATT	4050
GTCATTTAGG	AGAAAGGTCT	AAAGCIGGGC	TGAATGOCAT	GCACICATAG	4100
TCCCAGCTAC	TIGGGAGGCC	GACGIGAGAG	GATTGCTTGA	GICCIGGAGT	4150
TCAAGCCCAG	CCTGGGAAAC	ACAGIGAGAC	CICATIGCIA	TAAATAAATA	4200
AAATGAATAA	ATAAATAAAC	ACATAAATAA	ATTCATTAAA	TAAATAAAGT	4250
TITCATGGTA	TAGGAAAACA	CAGATGCAAA	GITTITIGICC	CIAGIGGCIG	4300
GIAATGITGC	AAACGTAACT	CCTTAGTGAA	CIGIACCACT	TAAAAATAGT	4350
	AATTITAGGA				4400
	TAAAGITCAG				4450
	TCTTCAAGAC				4500
	TCATATACIG	•			4550
	GTTTCACATC				4600
<b>4-1-</b>	ACACACACTC				4650
	AATTGTCTCG				4700
	AGGGIGIGIA				<b>475</b> 0
	AGCTOGAGGA				4800
	GCAGGICAAG				4850
	GIGIGIGIGI				4900
				ATGGAAAGAA	4950
	TGACTGGAAA				5000
				ATGAATGAGA	5050
				GACTGGTGAA	5100
				TCATGUATGT	5150
100104112	AAGGAAGTGA				5200
				CCAGTTCAGA	<b>525</b> 0
				COCCOCCITY	5300
				TOGATAAATG	5350
			•	CAGOGGOCIT	5400
				ACCGAATTAT	5450
				AAATGITTIG	5500
AAAGGAAGCA	TTIATGIAAC	GGAGGTCATG	GATTATTCCA	GGGATGCACT	5550

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GTTAAAAGTT	CCTAGAATCT	GACTGACAAC	AATGCCCATT	AATTGCTGTC	5600
CCCCACTCC	CTTATICICA	GTGCGGGGGA	CAGIATATIT	TCIGICATTC	5650
ACAAACAATG	TTATATTICG	TECTTTETTE	CTTCACGGGG	TICATITATIC	5700
GAATATTACC	TTTAGGACCT	TOGGACCIAA	ATATAACTTT	ATTICAACAA	5750
AGTGGAAGTT	TCTCTTTACC	CCGATAGGIA	ATCCCTCTCC	TGACTGTAAG	5800
ATTTOCATAG	TOCTCAAATC	CATCCAGCTA	ATCAATCCTT	CAGACOCTIGA	5850
CATIGIAATT	GIAACIGAAA	TOCTACOCAC	GTGGTAGACT	TCAGATTTCT	5900
CAGCTGACAC	ACACIGCIGI	TGGIACICIA	GGGCTGAATA	TAAGCATTAT	5950
ACATGTOCIG	TGGTTTATCC	TIAGAITGIC	ATTTAGGAGA	AAGGTCTAAA	6000
GCTGGGCTGA	ATGCCATGCA	CTCATAGICC	CAGCIACITG	GGAGGCCGAG	6050
GIGAGAGGAT	TGCTTGAGTC	CIGGAGITCA	AAGCCCAGCC	TGGGAAACAC	6100
AGTGAGACCT	CATTCCTAAT	AATAAATAA	ATGAATAAAT	AAATAAACAC	6150
TAAATAAAT	TCATTAAATA	AATAAAGITT	TCATGGTATA	GGAAAACACA	6200
GATGCAAAGT	TITICICCT	AGIGGCIGGI	AATGTTGCAA	ACGIAACTCC	6250
TIAGIGAACT	GIACCACITN	NNNTAGTTA	AGATOGTAAA	TITTAGGATA	6300
TCIGIATITT	TTACCACAAT	TGGAAATTOC	TTTCTTCCTA	AAGITCAGIG	6350
CAGITATCAT	ATATICTITT	AAATTTTTAC	TGIATGIATC	TICAAGACAT	6400
AACATICATA	GAAAATTTGC	AAGAATAGTA	CAATGAACTC	ATATACIGIT	6450
CATCIGGATT	CACCAAIGIG	GITAGTAGCT	TTOGCTTCAT	AGGITTCACA	6500
TCTTCTTCCC	TOGICICIT	ACCGIGCIGC	CCACACACIC	ACACACACAC	6550
ACTCACACAC	ACATACGGAT	ATATGTTTAC	TGITATTAAT	GIGAATIGIC	6600
TOGATAAAGT	TICAGGGATT	AIGGICCTIT	ACCCIATGIA	CITGAGGGIG	6650
TGIATATOGT	CAGAACAAAG	AGAAAGTCAT	TICTIGGATC	ATCACTGCAC	6700
AAAGATAAAA	ATCAGGAAAT	TIAACAATGA	GAAAATGGAG	TCATTTAATC	6750
ACAGAGIGCA	TACICAAATT	TIGOCAGCIT	CCCCAGAAAT	TICITITIC	6800
CTTTTTTTT	TCTTIGTICG	AGACGGAGIC	TCTCTCTGTG	GCCCAGGTTG	6850
GAGGGCAGTA	GIGOGAICIC	GGCICACIGC	AACCTACACC	TOOCAGGITC	6900
TAGGGATTCT	CATGCCTCAG	Cloccici	AGCIGGGACT	ACAGGOGGOG	6950
GCCACTGCGG	TCITGAACIT	CIGGCCICAC	CICCICICCC	CACCTIGGCA	7000
TCCCAAAATG	TTTGGATTGC .	AGGCGIGAGA	CCCACCACC	GGCCCAGATA	7050
ATTITATICA	TAGGATTICI	TTTTCTGATC	CAGAGICCAG	TIGAGAATCA	7100
CACCITCCAT	GIGCITITCA	GGIGITITA	GITICCTITA	ACCIGIAATG	7150
TTTCCTTAAT	TTTTCTTGIC	ATICACGATA	CCCACATTTT	TGGAGAGGAT	7200
AGACCAGTTG	GITICAGAA '	TATICIGIAG	TITGGGCTIT	TICATGUATT	7250
TTAAAAGAGT	TTICICACIC	AGOGTITATT	GGTGGCTACT	CATCCCATGT	7300
AAGAGICIAA	GUGUTAGGAG '	IGIAAGIGCT	GTGAGAGACG	GGATTTGAGC	7350
CTTGAGTCAT	ITAATACGAG	AAGGACAATC	AGAAGTAGAA	TAAGAGAGAA	7400

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GTGCAAAGGA	GGCAGCAAAG	TIGICIGAGG	<b>GCAGTCTTCG</b>	GAAAGGAGGA	7450
GGGINATATT	TGGAACACCT	TGITTICCIG	TITICICCIA	ACCGACTOCT	7500
GAAATAATGT	TOCTOGGATT	CITATCAACA	CATTTATTAT	TACGITAGCT	7550
AAAGCTTTTA	TATAATAATA	COGAGAGCAT	GAATATCATT	TICTIATICA	7600
TATTITATGT	TITACIGCIT	AAATTGATAC	GIATITITA	TTTTTTAAGGG	7650
COGAAGOCTG	AAGCTCATAG	CCAGGAACAG	GGTCACCCAC	AGACTGGGTG	7700
TGAGTGTGAA	GATGGTCCTG	ATGGGCAGGA	GATGGACCCG	CCAAATCCAG	<b>775</b> 0
AGGAGGTGAA	AACGCCTGAA	GAAGGTAGGC	AATCCATTAG	GCATGCACAT	7800
TGTAGGGTGT	CIGITICCAC	AGTATCATAT	TGTAACICIT	ACTATGTTTT	7850
TGAGACGGAG	TCTCCCTCTG	AAGACCAGGC	TGGAGTGCAG	TGGTGCCATC	7900
TOGGCTCACT	GGAAATTCTG	TCTCCAGGGT	TCAAGIGATT	CICCICCIC	<b>795</b> 0
AGCCTCTGGC	GGAGCCGGGC	TTACAGGCAT	CCICCCCCCC	GCCCAGCTAA	8000
TIGIIGIATT	TTTAGTAGAG	ACAGGGTTTC	GITATGTTGC	ACAGGITGIT	8050
CCCGAACTCC	TGACCICAGG	TGATOCACCT	GOCTOGACCA	TTGAAATTGC	8100
CGGGATTACA	GGCAGAGOCA	CCCTICCCCCA	CCCAGCATTA	TATTTTTAT	8150
AACAGAGAGG	TAACAATACT	GOGTCTTTAG	TAACAGAGTT	CITATATAAA	8200
GGTTATTTGA	AACGIAGIIC	AGGCCCCAGC	ACCOGGCIGA	TAGACTGTCA	8250
GATAGGGAAA	CAAAGIGAGI	CAAAGCTATG	TIGAATTAAA	AGITTICAGT	<b>830</b> 0
ATAAATCCIT	AAACCAGTAG	CICACAATTT	TCAGATGCTT	TIGIAAAGGT	8350
CIGCITITAA	TCAATACATA	ACACGITIGI	AACACOCATC	ACTIGGIGIG	8400
			ATACCAGCIC		8450
GGOGAGATIC	TGAGTGAGTC	CITICOCITC	TAAACCTATC	TITICGITCIT	8500
			TAAAACCATT	· · · · · · · · · · · · · · · · · · ·	8550
			AATGCGTGGG		8600
CATCICITOG	CATGGITIGI	TAATCACAAA	TCATGAAACC	CIGGOCCAG	8650
			TIGICAGIAA		<b>870</b> 0
			GTGGAGAGCT		8750
			CICITCITGG		8800
			AGGCTGGAGT		8850
			AGCTAAAGCC		8900
			CACAAGCCAC		8950
			TGAGGICICA		9000
			AATTGCTTCT		9050
			<b>GGIACATGIG</b>		9100
			CIGIGOGCIG		9150
			GIGCIATOOC		9200
CCCCACCCCA	CAACAGTOOC	CAGGGIGIGA	TATICCCCTT	CCICIGICCA	9250

46
3.3 updated LK new Sequence

	<del></del>				
10	20	30	40	50	
		1234567890			
		CCCACCTATG			9300
		GTTTACTGAG			9350
		ATGAACICIT			9400
TATTCCATAG	TGTATATGTG	CCACATTTTC	TTAATCCAGT	CTATOGITGT	9450
TGGACATTTG	GGITGGITCC	AAGICITIGC	TATOGIGAAT	AATGCCGCAA	9500
TAAACATAOG	TGTGCACGTG	TCTTTATAGC	AGCATGATTT	ATAGTCCTTT	9550
GGGTATATAC	CCAGTAATGG	GATGGCTGGG	TCAAATGGTA	CAATIGCITC	9600
TTAAATCTTT	CCCCACGGAA	ACCITGAGIG	ACTGAAATAA	ATATCAAATG	9650
GCGAGAGACC	GITTAGITCG	TATCATCIGT	GGCATGTAGG	TCAGTGATGC	9700
TCAGCATGGG	TGTGAGTAAG	ATGCCTGTGC	TATGCATGCT	CCCTGCCCCA	9750
		CIATTICIAA			9800
		ATATCAAATG			9850
		ATTTAAAGTT			9900
		AGTATGCTTT			9950
GCTGACTGTT	AATIGIAAGI	GCTTGTGACT	TGAAAGGAAG	CACATGATGT	10000
		AAATTCTGCA			10050
		<b>AAGACACGGG</b>			10100
		CCAAACIGIA			10150
GGGICATCIG	AATAATCIGT	TGCTAAGTCA	TGTTCCCAAT	CEFFEIGHT	10200
		TCACAGIGIT			10250
		TGGAAATTTG			10300
		AAAGAAGICT			10350
TATTICIAGC	TTTTTGATGC	TGTGAAATAT	GIATCATICT	TIGAAATOGT	10400
		GGTATGTAGA			10450
		TTTGAGACGG			10500
		TCTCTGCTCA			10550
TICAAGCAAT	TEICIECCIC	AGCCTCCCGA	GTAGCTGGGA	TTATAGGCAC	10600
		AGITITIGIGT			10650
CATCITIGGCC	GGGGIGCICI	TGAACICCIG	ACCICGIGAT	TCACCIGCCT	10700
TGGCCTCCCA	AAGIGCIGGG	ATTACAGGCA	TGCACGCCTC	CCCCCCCCCT	10750
GGAGACATAA	TICITACATA	TIGGITTICT	ATCCAGOGGC	CITCICAAAT	10800
ATGCTTGTGA	ATTCTAAAGT	TIACITCIAG	GICGITITCA	GICITCAATA	10850
TACAGAAACA	TATCATCCTG	GAATAAGAGC	AGITITGITT	CCCCATTIT	10900
TITTIGITIT	TCCTTTTGTA	CITITITIGE	AGAGACCCCC	TITIGCCATG	10950
TTTCCCCGGC	TGITGITGNN	NITITGAGIG	CAAGIGATGC	ACCCACGICA	11000
OCTOOCACAG	TGCTGGGATT	ACTOGOGTOG	GOCAGGGGCC	ACCOGIGGGG	11050
GCCCCCCICC	TICCCATICI	AAAGAGITITI	ATTICCITTT	CIGATITIAT	11100

47
3.3 updated LK new Sequence

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GGCATTGCGC	AGACCCACCC	GITACAATGG	TGACAGTGGA	CATCCTTGTC	11150
TTATCCCTGA	TGAGAAACCG	AAAAATTTCA	ACATTICCCC	ATOCIATICA	11200
CTCTCCTTTT	TITGTAGACG	GACTITATCA	GAGTGAGTCA	TIGCATICIG	11250
TTCCAAATTT	GCTGAGAGTA	TICATITGAA	TATATGITGA	TTTTCATCAA	11300
ACAGIGCAIC	TATTICGATT	ACCACAGOGT	TITITOCCAT	TCATGGGTTA	11350
ATATAGTGAA	TICGATIGAT	AAATTTGTAC	GITITIAGGT	TOGATTATTA	11400
AAACTIGAGA	CAGOGICICA	CICIGICACC	GAGGCTGGAG	TECCETCETC	11450
TTATCAGAGC	TC				11462

#### We claim:

- 1. Isolated nucleic acid molecule which encodes a GAGE tumor rejection antigen precursor, the complementary sequence of which hybridizes to SEQ ID NO: 29 under stringent conditions.
- The isolated nucleic acid molecule of claim 1, consisting of SEQ ID NO: 29.
- 3. Expression vector comprising the isolated nucleic acid molecule of claim 1, operably linked to a promoter.
- 4. Expression vector comprising the isolated nucleic acid molecule of claim 2, operably linked to a promoter.
- 5. Isolated eukaryotic cell transformed or transfected with the expression vector of claim 3.
- 6. Isolated eukaryotic cell transformed or transfected with the expression vector of claim 4.
- 7. Process for making an expression vector capable of encoding a GAGE tumor rejection antigen precursor, comprising inserting the isolated nucleic acid molecule of claim 1 into a vector which comprises a promoter, wherein said isolated nucleic acid molecule is inserted into said expression vector in operable linkage orienta-

49

tion to said promoter.

8. Process for making an expression vector capable of encoding a GAGE tumor rejection antigen precursor, comprising inserting the isolated nucleic acid molecule of claim 2 into a vector which comprises a promoter, wherein said isolated nucleic acid molecule is inserted into said expression vector in operable linkage orientation to said promoter.

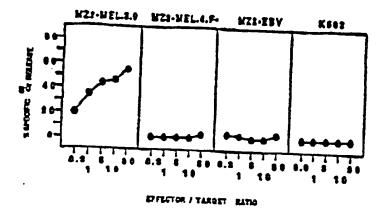
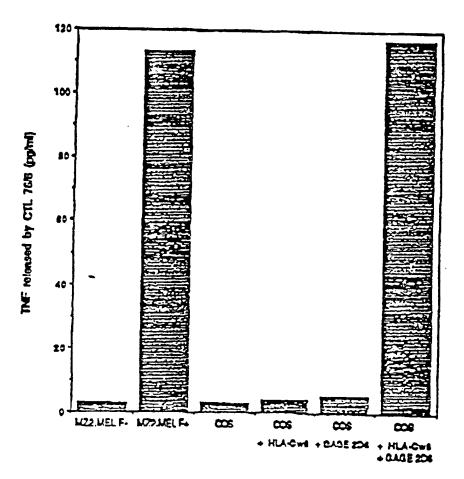


Figure 1



Norte 2

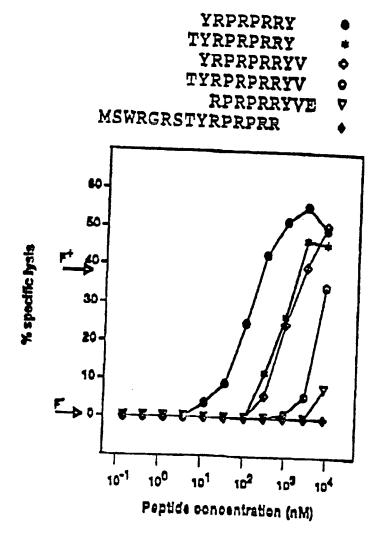
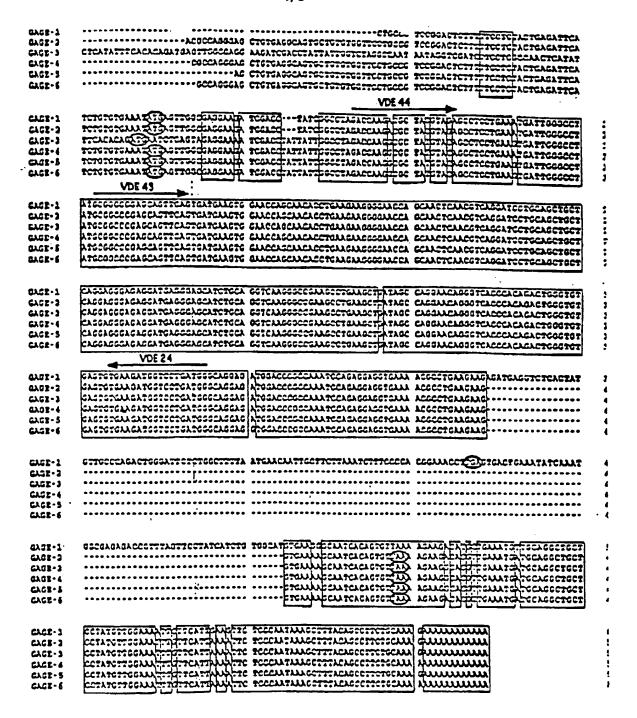
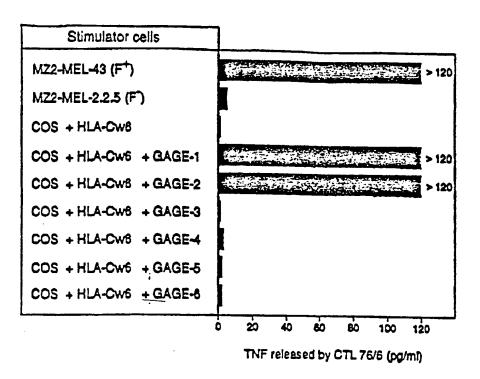


Figure 3



#### 5/8

	antigenic peptide		
GAGZ-1 GAGZ-2 GAGZ-3 GAGZ-4 GAGZ-5 GAGZ-6	# ####################################	GPMRPIQIEDEVIPATPEEGIFATO RODFALAGEGIDEGABAGOGPEPEA GPMRPIQIEDEVEPATPEEGIFATO RODFALAGIGEDEGASAGOGPEPEA GPMRPIQIEDEVEPATPEEGIFATO RODFALAGEGIDEGASAGOGPEPEA GPMRPIQIEDEVEPATPEEGIFATO RODFALAGEGIDEGASAGOGPEPEA GPMRPIQIEDEVEPATPEEGIFATO RODFALAGEGIDEGASAGOGPEPEA GPMRPIQIEDEVEPATPEEGIFATO RODFALAGEGIDEGASAGOGPEPEA	7: 7: 7: 7: 7:
Gage-1		PRPETVKTPELENGERTVAGTGILW LLDONGPLELEPREP	13
CYCE-5	EPOSOCKPOTCCCCCCCCOSPOS	PMPEEVK:PRICIKOROC	ii
Cace-3	DEGEOGRAPOTOTECHEOFPOGREPP	PMPERVATPERCEXOSOC	11
CYCE-1	DEGREES OF CONTROL OF	PRPERVRIPETCIKOSOC	11
GAGE-5	ಕರ್ಧಿವಿಕ್ಕರ್ಯವಿಕ್ಕಾರಿಕೆ ಕ್ರಮ್ ಕರ್ಮಿಕ್ಕರ್ಗಳಲ್ಲಿ	PRPIEVATPEDDINGEQC	11
GAGE-6	DESTANGE OF CONTROL OF THE PROPERTY OF	PHPLIVITPIDITEOSOC	11



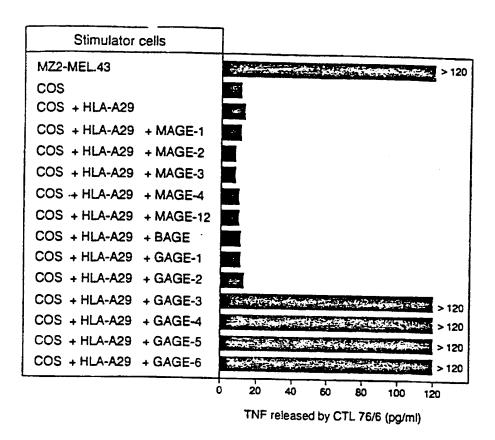


Figure 7. Stimulation of MZ2-CTL 22/23 by COS-7 cells transiently transfected with an HLA-A29 cDNA and MAGE, BAGE or GAGE cDNA. The CTL was added after 24 hours and the production of TNF was estimated 24 hours later. MZ2-MEL 43 was used as a positive control stimulator cell.

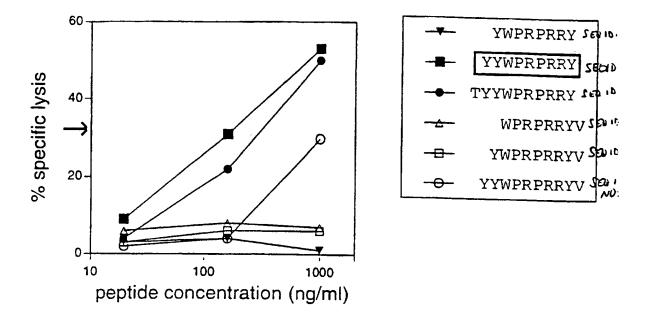


Figure §. Lysis by MZ2-CTL 22/23 of lymphoblastoid cell line LB17-EBV incubated with GAGE-encoded peptide YYWPRPRRY. Thousand 51Cr-labelled LB17-EBV target cells were incubated in 96 well microplates in the presence of various concentrations of peptide for 15 minutes at 37°C. An equal volume containing 6000 CTL was then added. Chromium release was measured after 4 hours at 37°C. We have indicated the final concentration of peptides during the incubation of the target cells with the CTL. The arrow indicates the percentage of lysis of MZ2-MEL.43 cells.



#### INTERNATIONAL SEARCH REPORT

International application No. PCT/US97/10850

			l				
	ICATION OF SUBJECT MATTER						
IPC(6) :A61K 38/00, 45/05, C07K 7/00, 14/82; C12N 15/00 US CL :536/23.5; 435/69.3, 320.1, 325							
According to International Patent Classification (IPC) or to both national classification and IPC							
	EARCHED						
	entation searched (classification system follow	ved by classification syr	nbols)	•			
U.S. : 536/2:	3.5; 435/69.3, 320.1, 325						
Documentation so	arched other than minimum documentation to	the extent that such docu	ments are include	d in the fields searched			
TUMOR REJE	CTION ANTIGEN FILES.						
Electronic data ba	se consulted during the international search (	name of data base and,	where practicable	c, search terms used)			
APS AND DIA	LOG (FILE-BIOCHEM) DATABASES: KE	Y WORDS: GAGE, TU	IMOR REJECTIO	ON ANTIGEN, DNA			
C. DOCUME	NTS CONSIDERED TO BE RELEVANT						
Category* C	itation of document, with indication, where a	ppropriate, of the releva	nt passages	Relevant to claim No.			
Y WO	) 95/03422 A1 (LUDWIG IN SEARCH) 02 February 1995, see	ISTITUTE FOR entire document.	CANCER	1-8			
VAN DEN EYNDE, B. et al. Presence on a human melanoma of multiple antigens recognized by autologous CTL. Intl J. Cancer. 1989, Vol. 44, pages 634-640, see entire document.				1-8			
WOLFEL, T. et al. Lysis of human melanoma cells by autologous cytolytic T cell clones. J. Exp. Med. September 1989, Vol. 170, pages 797-810, see entire document.				1-8			
Further docu	ments are listed in the continuation of Box C	See patent	family annex.				
	cories of cited documents:	*T* later document p	ublished efter the inte	rostional filing data or priority cation but cited to understand			
to be of part	fining the general state of the art which is not considered ioniar relevance	the principle or t	heory underlying the	invention			
	nent published on or after the international filing date	considered novel	or cannot be consider	ed to involve an inventive step			
cheed to estim	nich may throw doubts on priority claim(s) or which is blish the publication data of another citation or other o (as specified)		ont is taken elone ticular calavanca: the	chinad investing source by			
special rescon (as specified)  "Y"  document of perticular relevance; the claimed invention cannot be considered to involve an investive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art							
document published prior to the international filing data but later than "A" document member of the same patent family							
ate of the actual c	ate of the actual completion of the international search  Date of mailing of the international search report						
27 OCTOBER 19	97	0 5 DEC 199					
ame and mailing a Commissioner of Par	address of the ISA/US tents and Trademarks	Authorized officer					
Box PCT Washington, D.C. 2		THOMAS CUNNII	ngham ()	sab			
	3) 305-3230		309.0106	In			

Form PCT/ISA/210 (second sheet)(July 1992)\*

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

#### BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

□ BLACK BORDERS
□ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
□ FADED TEXT OR DRAWING
□ BLURRED OR ILLEGIBLE TEXT OR DRAWING
□ SKEWED/SLANTED IMAGES
□ COLOR OR BLACK AND WHITE PHOTOGRAPHS
□ GRAY SCALE DOCUMENTS
□ LINES OR MARKS ON ORIGINAL DOCUMENT
□ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

#### IMAGES ARE BEST AVAILABLE COPY.

☐ OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.